



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 17654

TO: Stephen Kapushoc
Location: 3a60 / 2c70
Tuesday, January 17, 2006
Art Unit: 1634
Phone: 571-272-3312
Serial Number: 10 / 754446

From: Jan Delaval
Location: Biotech-Chem Library
Remsen 1a51
Phone: 571-272-2504

jan.delaval@uspto.gov

Search Notes

THIS PAGE BLANK (USPTO)

STIC-Biotech/ChemLib

176 513

ME

From: Kapushoc, Stephen T.
Sent: Thursday, January 12, 2006 11:15 AM
To: STIC-Biotech/ChemLib
Subject: Sequence searches for 10/754,446

RECEIVED
JAN 12 2006
STIC

Please perform the following searches for application 10/754,446:

Score/Length sequence search for the following fragments of GenBank Locus AF287270, with a minimum hit length of 15 and a maximum hit length of 30, and a score over length cutoff at 90%:
Fragments of AF287270: 100-500, 6956-7356

Please also search the following sequences: SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7

441-460 = 1-20 SEQ 3

Thanks,
Steve

Stephen Kapushoc
Art Unit 1634 - USPTO
Tel: 571-272-3312
Office: REM 3A60
Mailbox: REM 2C70

09/851494

Goldin + Scierno

Searcher: Car
Searcher Phone: 22504
Date Searcher Picked up: 1/12/06
Date completed: 1/17/06
Searcher Prep Time: 10
Online Time: (50)

Type of Search
NA# ✓ AA#
S/L: c Oligomer:
Encode/Transl:
Structure #: Text:
Inventor: Litigation:

Vendors and cost where applicable
STN:
DIALOG:
QUESTEL/ORBIT:
LEXIS/NEXIS:
SEQUENCE SYSTEM: ✓
WWW/Internet:
Other (Specify):

THIS PAGE BLANK (USPTO)
THIS PAGE BLANK (USPTO)

sun + planet

c 107	17.4	4.3	19	1	ADY01278	Extend primer 154
c 108	17.4	4.3	19	1	ADX86792	XTAP targeting sir
c 109	17.4	4.3	19	1	ADX86325	XTAP targeting sir
c 110	17.4	4.3	19	1	ABE05189	Human IL-13RA1 sir
c 111	17.4	4.3	19	1	ABE05166	Human IL-13RA1 sir
c 112	17.4	4.3	19	1	ABE15226	Human IL-13RA1 tra
c 113	17.4	4.3	19	1	ABE15003	Human IL-13RA1 tra
c 114	17.2	4.3	19	1	AAQ76248	Generic primer fro
c 115	17.2	4.2	17	1	AAV23284	Nucleotide sequenc
c 116	17	4.2	17	1	AAC87597	Human Alu sequence
c 117	17	4.2	17	1	ADB44570	Tumour suppression
c 118	17	4.2	17	1	ACC51496	Human tumour suppr
c 119	17	4.2	17	1	ADL50193	Human PKR substrat
c 120	17	4.2	17	1	ADL50731	Human PKR substrat
c 121	17	4.2	17	1	ADL49907	Human PKR substrat
c 122	17	4.2	17	1	ADL49908	Human PKR substrat
c 123	17	4.2	17	1	ADL49909	Human PKR substrat
c 124	17	4.2	17	1	ADL49906	Human PKR substrat
c 125	17	4.2	17	1	ADL50192	Human PKR substrat
c 126	17	4.2	17	1	ADL82338	Human ER+ breast c
c 127	17	4.2	17	1	ADP08690	Extend primer 27 u
c 128	17	4.2	17	1	ADY01178	Extend primer 54 u
c 129	17	4.2	17	1	ADY00434	Extend primer 2 us
c 130	17	4.2	17	1	ADX82001	Melanoma associate
c 131	17	4.2	17	1	ADX81510	Melanoma associate
c 132	17	4.2	18	1	ADO56979	Human CARK/FPGT pr
c 133	17	4.2	18	1	ADO56561	Human cyclin-depen
c 134	16.4	4.1	18	1	AAZ27769	PCR primer for hum
c 135	16.4	4.1	18	1	ADY02910	Extend primer 360
c 136	16.4	4.1	18	1	ADY00671	Extend primer 239
c 137	16.4	4.1	18	1	ADY01277	Extend primer 153
c 138	16	4.0	16	1	AAF88161	Human thyroid malf
c 139	16	4.0	16	1	ACA62885	Repeated nucleic a
c 140	16	4.0	16	1	ACA62882	Repeated nucleic a
c 141	16	4.0	17	1	ADB04313	Human MD27 scannin
c 142	16	4.0	17	1	ADB04314	Human MD27 scannin
c 143	16	4.0	17	1	ACC63031	Murine oligonucleo
c 144	16	4.0	17	1	ADB44260	Tumour suppression
c 145	16	4.0	17	1	ACC51495	Human tumour suppr
c 146	16	4.0	17	1	ADL50732	Human PKR substrat
c 147	16	4.0	17	1	ADL49420	Human PKR substrat
c 148	16	4.0	17	1	ADX81033	Melanoma associate
c 149	15.4	3.8	17	1	AAA22718	Integrin subunit b
c 150	15.4	3.8	17	1	AAA22717	Integrin subunit b
c 151	15.4	3.8	17	1	ABT39409	Tumour suppression
c 152	15.4	3.8	17	1	ABT38213	Tumour suppression
c 153	15.4	3.8	17	1	ABT38720	Tumour suppression
c 154	15.4	3.8	17	1	ABT38728	Tumour suppression
c 155	15.4	3.8	17	1	ABT35457	Tumour suppression
c 156	15.4	3.8	17	1	ABT40068	Tumour suppression
c 157	15.4	3.8	17	1	ABT40150	Tumour suppression
c 158	15.4	3.8	17	1	ABT35874	Tumour suppression
c 159	15.4	3.8	17	1	ABZ60567	Human K-Ras DNazym
c 160	15.4	3.8	17	1	ADB41143	Tumour suppression
c 161	15.4	3.8	17	1	ADB43523	Tumour suppression
c 162	15.4	3.8	17	1	ADB44518	Tumour suppression
c 163	15.4	3.8	17	1	ADE43565	Human IDE sequenci
c 164	15.4	3.8	17	1	ADI50915	Human tumour suppr
c 165	15.4	3.8	17	1	ADI50723	Human tumour suppr
c 166	15.4	3.8	17	1	ADI52180	Human tumour suppr
c 167	15.4	3.8	17	1	ADI50051	Human tumour suppr
c 168	15.4	3.8	17	1	ADI51643	Human tumour suppr
c 169	15.4	3.8	17	1	ACC52610	Human tumour suppr
c 170	15.4	3.8	17	1	ACC51497	Human tumour suppr
c 171	15.4	3.8	17	1	ACC52025	Human tumour suppr
c 172	15.4	3.8	17	1	ACC53324	Human tumour suppr
c 173	15.4	3.8	17	1	ADL49905	Human PKR substrat
c 174	15.4	3.8	17	1	ADL49419	Human PKR substrat
c 175	15.4	3.8	17	1	ADL49904	Human PKR substrat
c 176	15.4	3.8	17	1	ADL50191	Human PKR substrat
c 177	15.4	3.8	17	1	ADL50733	Human PKR substrat
c 178	15.4	3.8	17	1	ADH54043	Human neurodegener
c 179	15.4	3.8	17	1	ADK13186	Human glioma endot
c 180	15.4	3.8	17	1	ADL82349	Human ER+ breast c
c 181	15.4	3.8	17	1	ADZ29641	Human K-Ras subestr
c 182	15.4	3.8	17	1	ABE71294	Extension primer f
c 183	15	3.7	15	1	AAF69438	Human IL4Ralpha ge
c 184	14	3.5	15	1	ABK98169	Triple helix formi
c 185	14	3.5	15	1	ABK98187	Triple helix formi
c 186	14	3.5	15	1	ABK98168	Triple helix formi
c 187	14	3.5	15	1	ABK98167	Triple helix formi
c 188	14	3.5	15	1	ABK98186	Triple helix formi
c 189	13.6	3.4	15	1	AAS98701	Colony stimulating
ALIGNMENTS						
RESULT 1						
ADB84281						
ID	ADB84281 standard; DNA; 25 BP.					
XX	AC ADB84281;					
XX	DT 04-DEC-2003 (first entry)					
XX	Human mucolipin (MCOLIPIN-1) gene AJ haplotype analysis primer #3.					
XX	human; mucolipin; MCOLN-1; mucopolipidosis IV; ion channel defect;					
KW	immunogen; gene therapy; ss; primer; PCR; AJ haplotype.					
XX	Homo sapiens.					
XX	US2003064363-A1.					
XX	03-APR-2003.					
PF	08-MAY-2001; 2001US-00851494.					
XX	18-AUG-2000; 2000US-0226388P.					
PR	(MLFO-) ML4 FOUND & HAVARD COLLEGE.					
PA	Goldin E, Slaugenhaupt SA, Sun M, Acierno JS;					
PI	WPI; 2003-540779/51.					
XX	New mucolipin (MCOLN1) polypeptides and nucleic acids, useful in gene					
PT	therapy for treating mucopolipidosis IV or ion channel defect, or as					
PT	immunogen to generate antibodies that recognize the MCOLN1 polypeptide.					
XX	Example; Page 18; 34pp; English.					
XX	The invention describes a nucleic acid (I) encoding a mucolipin (MCOLN1)					
CC	polypeptide, and sharing at least 95% sequence identity with sequence					
CC	selected from 2 nucleotide sequences (designated S1 and S2, not given in					
CC	the specification). A mutation of in the MCOLN1 gene encoding the MCOLN1					
CC	polypeptide results in a defect in expression of a functional MCOLN1. The					
CC	MCOLN1 gene is useful for the treatment of mucopolipidosis IV or					
CC	ion channel defect. The MCOLN1 polypeptides are useful as immunogen to					
CC	generate antibodies that recognize the MCOLN1 polypeptide. Such					
CC	antibodies may be used for diagnostics and intracellular regulation of					
CC	MCOLN1 activity. The polynucleotides may also be useful in gene therapy.					
CC	This sequence represents a primer used to analyse the AJ haplotype of the					
CC	human mucolipin (MCOLN-1) gene.					
XX	SQ Sequence 25 BP; 7 A; 5 C; 6 G; 7 T; 0 U; 0 Other;					
Query Match 6.2%; Score 25; DB 1; Length 25;						
Best Local Similarity 100.0%; Pred. No. 12;						
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
QY	226 CTGATATAAATGGCAGCGACTTTC 250					
DB	1 CTGATATAAATGGCAGCGACTTTC 25					

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2006, 14:56:42 ; Search time 2 Seconds
(without alignments)
2.069 Million cell updates/sec

Title: af287270

Perfect score: 401

Sequence: 1 GGGTCTCAGCGGGAACAA.....TAGCTCAGCGCTCGAC 401

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 0.5

Searched: 244 seqs, 5160 residues

Total number of hits satisfying chosen parameters: 488

Minimum DB seq length: 15

Maximum DB seq length: 30

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 244 summaries

Database : af287270_copy_100_500.rnpbm4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25.4	6.3	27	1	US-10-708-204-5106
2	25	6.2	25	1	US-10-956-157-173066
3	25	6.2	25	1	US-11-060-756-178198
4	24	6.0	24	1	US-10-708-204-5162
5	24	6.0	25	1	US-10-956-157-185762
6	24	6.0	25	1	US-10-956-157-208929
7	24	6.0	25	1	US-11-060-756-178197
8	24	6.0	25	1	US-11-060-756-203216
9	23.4	5.8	25	1	US-10-956-157-149516
10	23.4	5.8	25	1	US-10-956-157-149517
11	23.4	5.8	25	1	US-10-956-157-173133
12	23.4	5.8	25	1	US-10-956-157-177068
13	23.4	5.8	25	1	US-10-956-157-177069
14	23.4	5.8	25	1	US-10-956-157-184259
15	23.4	5.8	25	1	US-10-956-157-188388
16	23.4	5.8	25	1	US-10-956-157-273920
17	23.4	5.8	25	1	US-10-956-157-285780
18	23.4	5.8	25	1	US-10-956-157-286030
19	23.4	5.8	25	1	US-10-956-157-287740
20	23.4	5.8	25	1	US-10-956-157-293223
21	23.4	5.8	25	1	US-10-956-157-293224
22	23.4	5.8	25	1	US-10-956-157-293225
23	23.4	5.8	25	1	US-10-956-157-298488
24	23.4	5.8	25	1	US-11-060-756-204-5164
25	23.4	5.8	25	1	US-11-060-756-163912
26	23.4	5.8	25	1	US-11-060-756-164006
27	23.4	5.8	25	1	US-11-060-756-180494
28	23.4	5.8	25	1	US-11-060-756-200858
29	23.4	5.8	25	1	US-11-060-756-272590
30	23.4	5.8	25	1	US-11-060-756-274785
31	23.4	5.8	25	1	US-11-060-756-278943
32	23.4	5.8	25	1	US-11-060-756-279896
33	23.4	5.8	25	1	US-11-060-756-279897

23.4	5.8	25	1	US-11-060-756-287805	Sequence 287805,
23	5.7	23	1	US-09-851-494B-10	Sequence 10, Appl
23	5.7	23	1	US-10-708-204-5076	Sequence 5076, Ap
23	5.7	23	1	US-10-708-204-5099	Sequence 5099, Ap
22	5.5	22	1	US-10-708-204-7	Sequence 7, Appl
22	5.5	22	1	US-10-708-204-13	Sequence 13, Appl
22	5.5	22	1	US-10-708-204-72	Sequence 72, Appl
22	5.5	22	1	US-10-708-204-91	Sequence 91, Appl
22	5.5	22	1	US-10-708-204-195	Sequence 195, App
22	5.5	22	1	US-10-708-204-208	Sequence 208, App
22	5.5	22	1	US-10-708-204-771	Sequence 771, App
22	5.5	22	1	US-10-708-204-1122	Sequence 1122, App
22	5.5	22	1	US-10-708-204-1127	Sequence 1127, Ap
22	5.5	22	1	US-10-708-204-1177	Sequence 1177, Ap
22	5.5	22	1	US-10-708-204-1399	Sequence 1399, Ap
22	5.5	22	1	US-10-708-204-1401	Sequence 1401, Ap
22	5.5	22	1	US-10-708-204-1480	Sequence 1480, Ap
22	5.5	22	1	US-10-708-204-1498	Sequence 1498, Ap
22	5.5	22	1	US-10-708-204-2082	Sequence 2082, Ap
22	5.5	22	1	US-10-708-204-2092	Sequence 2092, Ap
22	5.5	22	1	US-10-708-204-2517	Sequence 2517, Ap
22	5.5	22	1	US-10-708-204-2540	Sequence 2540, Ap
22	5.5	22	1	US-10-708-204-301	Sequence 301, Ap
22	5.5	22	1	US-10-708-204-5075	Sequence 5075, Ap
22	5.5	22	1	US-10-708-204-5086	Sequence 5086, Ap
22	5.5	22	1	US-10-708-204-5095	Sequence 5095, Ap
22	5.5	22	1	US-10-708-204-5098	Sequence 5098, Ap
22	5.5	22	1	US-10-708-204-5105	Sequence 5105, Ap
22	5.5	22	1	US-10-708-204-3821	Sequence 3821, Ap
22	5.5	22	1	US-10-708-204-3863	Sequence 3863, Ap
22	5.5	22	1	US-10-708-204-5124	Sequence 5124, Ap
22	5.5	22	1	US-10-126-103-235	Sequence 235, App
22	5.5	22	1	US-10-431-096-235	Sequence 235, App
22	5.5	22	1	US-10-708-204-3851	Sequence 3851, Ap
22	5.5	22	1	US-10-708-204-3870	Sequence 3870, Ap
22	5.5	22	1	US-10-708-204-5096	Sequence 5096, Ap
22	5.5	22	1	US-10-708-204-5165	Sequence 5165, Ap
22	5.5	22	1	US-10-708-204-5097	Sequence 5097, Ap
22	5.5	22	1	US-09-918-686-93	Sequence 93, Appl
22	5.5	22	1	US-10-353-150-93	Sequence 93, Appl
22	5.5	22	1	US-10-374-077-11	Sequence 11, Appl
22	5.5	22	1	US-10-708-204-14	Sequence 14, Appl
22	5.5	22	1	US-10-708-204-73	Sequence 73, Appl
22	5.5	22	1	US-10-708-204-75	Sequence 75, Appl
22	5.5	22	1	US-10-708-204-162	Sequence 162, App
22	5.5	22	1	US-10-708-204-192	Sequence 192, App
22	5.5	22	1	US-10-708-204-193	Sequence 193, App
22	5.5	22	1	US-10-708-204-194	Sequence 194, App
22	5.5	22	1	US-10-708-204-458	Sequence 458, App
22	5.5	22	1	US-10-708-204-460	Sequence 460, App
22	5.5	22	1	US-10-708-204-474	Sequence 474, App
22	5.5	22	1	US-10-708-204-475	Sequence 475, App
22	5.5	22	1	US-10-708-204-666	Sequence 666, App
22	5.5	22	1	US-10-708-204-772	Sequence 772, App
22	5.5	22	1	US-10-708-204-872	Sequence 872, App
22	5.5	22	1	US-10-708-204-899	Sequence 899, App
22	5.5	22	1	US-10-708-204-1078	Sequence 1078, Ap
22	5.5	22	1	US-10-708-204-1107	Sequence 1107, Ap
22	5.5	22	1	US-10-708-204-1108	Sequence 1108, Ap
22	5.5	22	1	US-10-708-204-1109	Sequence 1109, Ap
22	5.5	22	1	US-10-708-204-1113	Sequence 1113, Ap
22	5.5	22	1	US-10-708-204-1400	Sequence 1400, Ap
22	5.5	22	1	US-10-708-204-1478	Sequence 1478, Ap
22	5.5	22	1	US-10-708-204-1479	Sequence 1479, Ap
22	5.5	22	1	US-10-708-204-1482	Sequence 1482, Ap
22	5.5	22	1	US-10-708-204-1548	Sequence 1548, Ap
22	5.5	22	1	US-10-708-204-1556	Sequence 1556, Ap
22	5.5	22	1	US-10-708-204-1557	Sequence 1557, Ap
22	5.5	22	1	US-10-708-204-1558	Sequence 1558, Ap
22	5.5	22	1	US-10-708-204-1587	Sequence 1587, Ap

c 107	20.4	5.1	22	1	US-10-708-204-1595	Sequence 1595, Ap	c 180	18.4	4.6	20	1	US-10-199-676-74	Sequence 74, Appl
c 108	20.4	5.1	22	1	US-10-708-204-1720	Sequence 1720, Ap	c 181	18.4	4.6	20	1	US-10-671-395-383	Sequence 383, App
c 109	20.4	5.1	22	1	US-10-708-204-1886	Sequence 1886, Ap	c 182	18.4	4.6	20	1	US-10-671-395-752	Sequence 752, App
c 110	20.4	5.1	22	1	US-10-708-204-2091	Sequence 2091, Ap	c 183	18.4	4.6	20	1	US-10-671-395-948	Sequence 948, App
c 111	20.4	5.1	22	1	US-10-708-204-2197	Sequence 2197, Ap	c 184	18.4	4.6	20	1	US-10-671-395-1041	Sequence 1041, Ap
c 112	20.4	5.1	22	1	US-10-708-204-2207	Sequence 2207, Ap	c 185	18.4	4.6	20	1	US-10-671-395-1101	Sequence 1101, Ap
c 113	20.4	5.1	22	1	US-10-708-204-2248	Sequence 2248, Ap	c 186	18.4	4.6	20	1	US-10-671-395-1167	Sequence 1167, Ap
c 114	20.4	5.1	22	1	US-10-708-204-2249	Sequence 2249, Ap	c 187	18.4	4.6	20	1	US-10-671-395-1231	Sequence 1231, Ap
c 115	20.4	5.1	22	1	US-10-708-204-2250	Sequence 2250, Ap	c 188	18.4	4.6	20	1	US-10-671-395-1323	Sequence 1323, Ap
c 116	20.4	5.1	22	1	US-10-708-204-2252	Sequence 2252, Ap	c 189	18.4	4.6	20	1	US-10-671-395-1549	Sequence 1549, Ap
c 117	20.4	5.1	22	1	US-10-708-204-2516	Sequence 2516, Ap	c 190	18.4	4.6	20	1	US-10-671-395-1567	Sequence 1567, Ap
c 118	20.4	5.1	22	1	US-10-708-204-2518	Sequence 2518, Ap	c 191	18.4	4.6	20	1	US-10-639-300-38	Sequence 38, Appl
c 119	20.4	5.1	22	1	US-10-708-204-3820	Sequence 3820, Ap	c 192	18.4	4.6	20	1	US-10-639-300-74	Sequence 74, Appl
c 120	20.4	5.1	22	1	US-10-708-204-3827	Sequence 3827, Ap	c 193	18.4	4.6	20	1	US-10-840-590-1509	Sequence 1509, Ap
c 121	20.4	5.1	22	1	US-10-708-204-3862	Sequence 3862, Ap	c 194	18.4	4.6	20	1	US-10-840-590-1510	Sequence 1510, Ap
c 122	20.4	5.1	22	1	US-10-708-204-5114	Sequence 5114, Ap	c 195	18.2	4.5	19	1	US-10-831-819-7	Sequence 7, Appli
c 123	20	5.0	20	1	US-10-085-906-302	Sequence 302, App	c 196	18	4.5	18	1	US-10-708-204-5101	Sequence 5101, Ap
c 124	20	5.0	20	1	US-10-251-699-1	Sequence 1, Appli	c 197	18	4.5	19	1	US-10-098-871-37	Sequence 37, Appl
c 125	20	5.0	20	1	US-10-172-911-80	Sequence 80, Appl	c 198	18	4.5	19	1	US-10-800-487-155	Sequence 155, App
c 126	20	5.0	20	1	US-10-189-268-71	Sequence 71, Appl	c 199	18	4.5	19	1	US-10-800-487-321	Sequence 321, App
c 127	20	5.0	20	1	US-10-754-446-3	Sequence 3, Appli	c 200	18	4.5	19	1	US-10-923-181-155	Sequence 155, App
c 128	20	5.0	20	1	US-10-708-204-3841	Sequence 3841, Ap	c 201	18	4.5	19	1	US-10-923-181-321	Sequence 321, App
c 129	20	5.0	20	1	US-10-708-204-5094	Sequence 5094, Ap	c 202	18	4.5	20	1	US-09-993-731-22	Sequence 22, Appl
c 130	20	5.0	21	1	US-10-708-204-3869	Sequence 3869, Ap	c 203	18	4.5	20	1	US-10-708-204-4383	Sequence 4383, Ap
c 131	20	5.0	21	1	US-10-708-204-5108	Sequence 5108, Ap	c 204	18	4.5	20	1	US-10-863-973-922	Sequence 922, App
c 132	20	5.0	22	1	US-10-708-204-95	Sequence 95, Appl	c 205	17.4	4.3	19	1	US-10-863-973-1145	Sequence 1145, Ap
c 133	20	5.0	22	1	US-10-708-204-202	Sequence 202, App	c 206	17.4	4.3	19	1	US-10-922-675-922	Sequence 922, App
c 134	20	5.0	22	1	US-10-708-204-382	Sequence 382, App	c 207	17.4	4.3	19	1	US-10-922-675-1145	Sequence 1145, Ap
c 135	20	5.0	22	1	US-10-708-204-483	Sequence 483, App	c 208	17.4	4.3	19	1	US-11-014-373-190	Sequence 190, App
c 136	20	5.0	22	1	US-10-708-204-831	Sequence 831, App	c 209	17.4	4.3	19	1	US-11-014-373-467	Sequence 467, App
c 137	20	5.0	22	1	US-10-708-204-876	Sequence 876, App	c 210	17.4	4.3	19	1	US-10-463-9818-2	Sequence 2, Appli
c 138	20	5.0	22	1	US-10-708-204-886	Sequence 886, App	c 211	17.2	4.3	18	1	US-09-242-772-1	Sequence 1, Appli
c 139	20	5.0	22	1	US-10-708-204-1047	Sequence 1047, Ap	c 212	17	4.2	17	1	US-10-156-306-1651	Sequence 1651, Ap
c 140	20	5.0	22	1	US-10-708-204-1102	Sequence 1102, Ap	c 213	17	4.2	17	1	US-10-156-306-1652	Sequence 1652, Ap
c 141	20	5.0	22	1	US-10-708-204-1123	Sequence 1123, Ap	c 214	17	4.2	17	1	US-10-156-306-1653	Sequence 1653, Ap
c 142	20	5.0	22	1	US-10-708-204-1176	Sequence 1176, Ap	c 215	17	4.2	17	1	US-10-156-306-1654	Sequence 1654, Ap
c 143	20	5.0	22	1	US-10-708-204-1487	Sequence 1487, Ap	c 216	17	4.2	17	1	US-10-156-306-2389	Sequence 2389, Ap
c 144	20	5.0	22	1	US-10-708-204-1539	Sequence 1539, Ap	c 217	17	4.2	17	1	US-10-156-306-2390	Sequence 2390, Ap
c 145	20	5.0	22	1	US-10-708-204-1653	Sequence 1653, Ap	c 218	17	4.2	17	1	US-10-156-306-3776	Sequence 3776, Ap
c 146	20	5.0	22	1	US-10-708-204-2243	Sequence 2243, Ap	c 219	17	4.2	17	1	US-10-339-782-309	Sequence 309, App
c 147	20	5.0	22	1	US-10-708-204-2274	Sequence 2274, Ap	c 220	17	4.2	17	1	US-10-704-513-302	Sequence 302, App
c 148	20	5.0	22	1	US-10-708-204-2548	Sequence 2548, Ap	c 221	17	4.2	17	1	US-10-704-513-720	Sequence 720, App
c 149	20	5.0	22	1	US-10-708-204-4312	Sequence 4312, Ap	c 222	17	4.2	17	1	US-10-840-590-793	Sequence 793, App
c 150	20	5.0	22	1	US-10-708-204-4381	Sequence 4381, Ap	c 223	17	4.2	17	1	US-10-840-590-1284	Sequence 1284, Ap
c 151	20	5.0	22	1	US-10-708-204-4385	Sequence 4385, Ap	c 224	17	4.2	17	1	US-09-784-423-69	Sequence 69, Appl
c 152	20	5.0	22	1	US-10-708-204-5072	Sequence 5072, Ap	c 225	16.4	4.1	18	1	US-10-708-204-3825	Sequence 3825, Ap
c 153	19.4	4.8	21	1	US-10-085-906-401	Sequence 401, App	c 226	16.4	4.1	18	1	US-09-739-909-8	Sequence 8, Appli
c 154	19.4	4.8	21	1	US-10-216-122-116	Sequence 116, App	c 227	16	4.0	16	1	US-09-739-909-11	Sequence 11, Appl
c 155	19.4	4.8	21	1	US-10-255-434-25	Sequence 25, Appl	c 228	16	4.0	16	1	US-10-156-306-534	Sequence 534, App
c 156	19.4	4.8	21	1	US-10-374-077-7	Sequence 7, Appli	c 229	16	4.0	17	1	US-10-156-306-3777	Sequence 3777, Ap
c 157	19.4	4.8	21	1	US-10-708-204-3826	Sequence 3826, Ap	c 230	16	4.0	17	1	US-10-703-817-224	Sequence 224, App
c 158	19	4.7	19	1	US-10-708-204-3840	Sequence 3840, Ap	c 231	16	4.0	17	1	US-10-840-590-314	Sequence 314, App
c 159	19	4.7	19	1	US-10-708-204-3842	Sequence 3842, Ap	c 232	16	4.0	17	1	US-10-156-306-533	Sequence 533, App
c 160	19	4.7	19	1	US-10-708-204-3854	Sequence 3854, Ap	c 233	15.4	3.8	17	1	US-10-156-306-1649	Sequence 1649, Ap
c 161	19	4.7	19	1	US-10-708-204-4379	Sequence 4379, Ap	c 234	15.4	3.8	17	1	US-10-156-306-1650	Sequence 1650, Ap
c 162	19	4.7	19	1	US-10-708-204-5110	Sequence 5110, Ap	c 235	15.4	3.8	17	1	US-10-156-306-2388	Sequence 2388, Ap
c 163	19	4.7	19	1	US-10-708-204-5172	Sequence 5172, Ap	c 236	15.4	3.8	17	1	US-10-156-306-3778	Sequence 3778, Ap
c 164	19	4.7	20	1	US-09-752-983-243	Sequence 243, App	c 237	15.4	3.8	17	1	US-10-438-700-679	Sequence 679, App
c 165	19	4.7	20	1	US-09-863-806-155	Sequence 155, App	c 238	15.4	3.8	17	1	US-10-339-782-320	Sequence 320, App
c 166	19	4.7	20	1	US-10-005-344-243	Sequence 243, App	c 239	15.4	3.8	17	1	US-10-339-793-16	Sequence 16, Appl
c 167	19	4.7	20	1	US-10-671-395-395	Sequence 395, App	c 240	15.4	3.8	17	1	US-10-282-174-170	Sequence 170, App
c 168	19	4.7	20	1	US-10-671-395-422	Sequence 422, App	c 241	15.4	3.8	17	1	US-10-600-009-170	Sequence 170, App
c 169	19	4.7	20	1	US-10-754-478-155	Sequence 155, App	c 242	15.4	3.8	17	1	US-10-724-270-679	Sequence 679, App
c 170	19	4.7	20	1	US-10-708-204-3868	Sequence 3868, Ap	c 243	15.4	3.8	17	1	US-10-010-802-81	Sequence 81, Appl
c 171	19	4.7	21	1	US-10-708-204-4384	Sequence 4384, Ap	c 244	15	3.7	15	1		
c 172	19	4.7	21	1	US-10-708-204-5102	Sequence 5102, Ap							
c 173	18.4	4.6	20	1	US-09-752-983-241	Sequence 241, App							
c 174	18.4	4.6	20	1	US-09-752-983-242	Sequence 242, App							
c 175	18.4	4.6	20	1	US-10-181-177-94	Sequence 94, Appl							
c 176	18.4	4.6	20	1	US-10-005-344-241	Sequence 241, App							
c 177	18.4	4.6	20	1	US-10-005-344-242	Sequence 242, App							
c 178	18.4	4.6	20	1	US-10-148-355A-64	Sequence 64, Appl							
c 179	18.4	4.6	20	1	US-10-199-676-38	Sequence 38, Appl							

ALIGNMENTS

RESULT 1
US-10-708-204-5106
; Sequence 5106, Application us/10708204

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2006, 14:58:13 ; Search time 0.001 Seconds
(without alignments)
318.394 Million cell updates/sec

Title: af287270
Perfect score: 401
Sequence: 1 GGGTCTCAGCAGGGAACAA.....TAGTCTACTGCGCTCGAC 401

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 0.5

Searched: 19 segs, 397 residues

Total number of hits satisfying chosen parameters: 38

Minimum DB seq length: 15
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 19 summaries

Database : af287270_copy_100_500.rnpbn4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	23.4	5.8	25	1 US-11-121-849-17198	Sequence 17198, A
C 2	23.4	5.8	25	1 US-11-121-849-44508	Sequence 44508, A
C 3	23.4	5.8	25	1 US-11-121-849-44509	Sequence 44509, A
C 4	23.4	5.8	25	1 US-11-121-849-48982	Sequence 48982, A
C 5	23.4	5.8	25	1 US-11-121-849-48983	Sequence 48983, A
C 6	23.4	5.8	25	1 US-11-121-849-48984	Sequence 48984, A
C 7	23.4	5.8	25	1 US-11-121-849-48985	Sequence 48985, A
C 8	20	5.0	20	1 US-10-515-538-80	Sequence 80, Appl
C 9	19	4.7	20	1 US-10-353-783-33	Sequence 33, Appl
C 10	18	4.5	18	1 US-10-857-780-2580	Sequence 2580, Ap
C 11	18	4.5	19	1 US-10-857-780-4587	Sequence 4587, Ap
C 12	17.4	4.3	19	1 US-10-857-780-2706	Sequence 2706, Ap
C 13	17.4	4.3	19	1 US-11-001-347-922	Sequence 922, App
C 14	17.4	4.3	19	1 US-11-001-347-1145	Sequence 1145, Ap
C 15	17	4.2	17	1 US-10-857-780-1868	Sequence 1868, Ap
C 16	17	4.2	17	1 US-10-857-780-2606	Sequence 2606, Ap
C 17	16.4	4.1	18	1 US-10-857-780-2105	Sequence 2105, Ap
C 18	16.4	4.1	18	1 US-10-857-780-2705	Sequence 2705, Ap
C 19	16.4	4.1	18	1 US-10-857-780-4341	Sequence 4341, Ap

ALIGNMENTS

RESULT 1
US-11-121-849-17198/c
; Sequence 17198, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
; APPLICANT: John Palma
; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
; FILE OF INVENTION: Microarrays
; FILE REFERENCE: 3684.1

; CURRENT APPLICATION NUMBER: US/11/121,849
; CURRENT FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: 60/567,949
; PRIOR FILING DATE: 2004-05-03
; NUMBER OF SEQ ID NOS: 673904
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 17198
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-11-121-849-17198

Query Match 5.8%; Score 23.4; DB 1; Length 25;
Best Local Similarity 96.0%; Pred. No. 2;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 440 TCTTGCTCTGTTGCCAGGCTGGAG 464
|||||
Db 25 TCTTGCTCTGTTGCCAGGCTGGAG 1

RESULT 2

US-11-121-849-44508/c
; Sequence 44508, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
; APPLICANT: John Palma
; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
; FILE REFERENCE: 3684.1
; CURRENT APPLICATION NUMBER: US/11/121,849
; CURRENT FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: 60/567,949
; PRIOR FILING DATE: 2004-05-03
; NUMBER OF SEQ ID NOS: 673904
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 44508
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-11-121-849-44508

Query Match 5.8%; Score 23.4; DB 1; Length 25;
Best Local Similarity 96.0%; Pred. No. 2;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 470 TGGTGTGATCATAGCTCACTGCAGC 494
|||||
Db 25 TGGTGTGATCATAGCTCACTGCAGC 1

RESULT 3

US-11-121-849-44509/c
; Sequence 44509, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
; APPLICANT: John Palma
; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
; FILE REFERENCE: 3684.1
; CURRENT APPLICATION NUMBER: US/11/121,849
; CURRENT FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: 60/567,949
; PRIOR FILING DATE: 2004-05-03
; NUMBER OF SEQ ID NOS: 673904
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 44509
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-11-121-849-44509

Query Match 5.8%; Score 23.4; DB 1; Length 25;

```
Best Local Similarity 96.0%; Pred. No. 2;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 465 TGCAGTGGTGCATAGTCTCACT 489
Db 25 TGCAGTGGTGTGATGATGCTCACT 1

RESULT 4
US-11-121-849-48982/c
; Sequence 48982, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
; APPLICANT: John Palma
; TITLE OF INVENTION: Microarrays
; FILE REFERENCE: 3684.1
; CURRENT APPLICATION NUMBER: US/11/121,849
; CURRENT FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: 60/567,949
; PRIOR FILING DATE: 2004-05-03
; NUMBER OF SEQ ID NOS: 673904
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 48982
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-11-121-849-48982

Query Match 5.8%; Score 23.4; DB 1; Length 25;
Best Local Similarity 96.0%; Pred. No. 2;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 447 CTGTGCCCCAGGCTGGAGTGCAGTG 471
Db 25 CTGTGCCCCAGGCTGGAGTGCAGTG 1

RESULT 5
US-11-121-849-48983/c
; Sequence 48983, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
; APPLICANT: John Palma
; TITLE OF INVENTION: Microarrays
; FILE REFERENCE: 3684.1
; CURRENT APPLICATION NUMBER: US/11/121,849
; CURRENT FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: 60/567,949
; PRIOR FILING DATE: 2004-05-03
; NUMBER OF SEQ ID NOS: 673904
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 48983
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-11-121-849-48983

Query Match 5.8%; Score 23.4; DB 1; Length 25;
Best Local Similarity 96.0%; Pred. No. 2;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 446 TCTGTGCCCCAGGCTGGAGTGCAT 470
Db 25 TCTGTGCCCCAGGCTGGAGTGCAT 1

RESULT 6
US-11-121-849-48984/c
; Sequence 48984, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
; APPLICANT: John Palma
; TITLE OF INVENTION: Microarrays
; FILE REFERENCE: 3684.1
; CURRENT APPLICATION NUMBER: US/11/121,849
; CURRENT FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: 60/567,949
; PRIOR FILING DATE: 2004-05-03
; NUMBER OF SEQ ID NOS: 673904
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 48984
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-11-121-849-48984

Query Match 5.8%; Score 23.4; DB 1; Length 25;
Best Local Similarity 96.0%; Pred. No. 2;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 445 CTCTGTGCCCCAGGCTGGAGTGCAG 469
Db 25 CTCTGTGCCCCAGGCTGGAGTGCAG 1

RESULT 7
US-11-121-849-48985/c
; Sequence 48985, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
; APPLICANT: John Palma
; TITLE OF INVENTION: Microarrays
; FILE REFERENCE: 3684.1
; CURRENT APPLICATION NUMBER: US/11/121,849
; CURRENT FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: 60/567,949
; PRIOR FILING DATE: 2004-05-03
; NUMBER OF SEQ ID NOS: 673904
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 48985
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-11-121-849-48985

Query Match 5.8%; Score 23.4; DB 1; Length 25;
Best Local Similarity 96.0%; Pred. No. 2;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 444 GCTCTGTGCCCCAGGCTGGAGTGCA 468
Db 25 GCTCTGTGCCCCAGGCTGGAGTGCA 1

RESULT 8
US-10-515-538-80
; Sequence 80, Application US/10515538
; Publication No. US20050282760A1
; GENERAL INFORMATION:
; APPLICANT: Isis Pharmaceuticals, Inc.
; APPLICANT: Lex M. Cowsett
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF PTPN12 EXPRESSION
; FILE REFERENCE: PFS-0016USA
; CURRENT APPLICATION NUMBER: US/10/515,538
; CURRENT FILING DATE: 2004-11-23
; PRIOR APPLICATION NUMBER: 10/172,911
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US03/18707
; PRIOR FILING DATE: 2003-06-12
; NUMBER OF SEQ ID NOS: 123
; SEQ ID NO 80
```

```
APPLICANT: John Palma
; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
; FILE REFERENCE: 3684.1
; CURRENT APPLICATION NUMBER: US/11/121,849
; CURRENT FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: 60/567,949
; PRIOR FILING DATE: 2004-05-03
; NUMBER OF SEQ ID NOS: 673904
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 48984
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-11-121-849-48984

Query Match 5.8%; Score 23.4; DB 1; Length 25;
Best Local Similarity 96.0%; Pred. No. 2;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 445 CTCTGTGCCCCAGGCTGGAGTGCAG 469
Db 25 CTCTGTGCCCCAGGCTGGAGTGCAG 1

RESULT 7
US-11-121-849-48985/c
; Sequence 48985, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
; APPLICANT: John Palma
; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
; FILE REFERENCE: 3684.1
; CURRENT APPLICATION NUMBER: US/11/121,849
; CURRENT FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: 60/567,949
; PRIOR FILING DATE: 2004-05-03
; NUMBER OF SEQ ID NOS: 673904
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 48985
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-11-121-849-48985

Query Match 5.8%; Score 23.4; DB 1; Length 25;
Best Local Similarity 96.0%; Pred. No. 2;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 444 GCTCTGTGCCCCAGGCTGGAGTGCA 468
Db 25 GCTCTGTGCCCCAGGCTGGAGTGCA 1

RESULT 8
US-10-515-538-80
; Sequence 80, Application US/10515538
; Publication No. US20050282760A1
; GENERAL INFORMATION:
; APPLICANT: Isis Pharmaceuticals, Inc.
; APPLICANT: Lex M. Cowsett
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF PTPN12 EXPRESSION
; FILE REFERENCE: PFS-0016USA
; CURRENT APPLICATION NUMBER: US/10/515,538
; CURRENT FILING DATE: 2004-11-23
; PRIOR APPLICATION NUMBER: 10/172,911
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US03/18707
; PRIOR FILING DATE: 2003-06-12
; NUMBER OF SEQ ID NOS: 123
; SEQ ID NO 80
```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2006, 14:54:56 ; Search time 0.001 Seconds
(without alignments)
874.180 Million cell updates/sec

Title: af287270
Perfect score: 401
Sequence: 1 GGGGTCTCAGCAGGACAA.....TAGCTCACTGCGCCTCGAC 401

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.5

Searched: 58 seqs, 1090 residues
Total number of hits satisfying chosen parameters: 116

Minimum DB seq length: 15
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 58 summaries

Database : af287270_copy_100_500.rni4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20.4	5.1	22	1	US-08-599-252-35
C 2	20.4	5.1	22	1	US-08-599-252-38
C 3	20.4	5.1	22	1	US-08-291-074-2
C 4	20.4	5.1	22	1	US-08-781-891-11
C 5	20.4	5.1	22	1	US-09-918-686-93
C 6	20.4	5.1	22	1	US-09-618-166-11
C 7	20.4	5.1	22	1	PCT-US96-06352-35
C 8	20.4	5.1	22	1	PCT-US96-06352-38
C 9	20.4	5.1	22	1	PCT-US96-06583-35
C 10	20.4	5.1	22	1	PCT-US96-06583-38
C 11	20.4	5.1	22	1	US-09-418-804-1
C 12	20.4	5.0	20	1	US-10-172-911-80
C 13	20.4	5.0	20	1	US-10-251-699-1
C 14	20.4	5.0	21	1	US-08-753-147-28
C 15	19.4	4.8	21	1	US-08-781-891-7
C 16	19.4	4.8	21	1	US-08-847-844A-116
C 17	19.4	4.8	21	1	US-09-618-166-7
C 18	19.4	4.7	20	1	US-09-280-805-243
C 19	19.4	4.7	20	1	US-09-038-637-155
C 20	19.4	4.7	20	1	US-09-224-683-33
C 21	19.4	4.7	20	1	US-09-604-325A-33
C 22	18.4	4.6	20	1	US-08-222-177A-274
C 23	18.4	4.6	20	1	US-09-280-805-241
C 24	18.4	4.6	20	1	US-09-280-805-242
C 25	18.4	4.6	20	1	US-09-487-445-94
C 26	18.4	4.6	20	1	US-09-467-642-64
C 27	18.4	4.6	20	1	US-09-780-173A-18
C 28	18.2	4.5	19	1	US-08-070-517-2
C 29	18.2	4.5	19	1	US-08-118-441-2
C 30	18.2	4.5	19	1	US-08-422-699A-14
C 31	18.2	4.5	19	1	US-08-422-706B-14
C 32	18.2	4.5	19	1	US-08-338-579A-2
C 33	18.2	4.5	19	1	PCT-US94-09851-2

34	17.4	4.3	19	1	US-09-366-840-2	Sequence 2, Appli
35	17.2	4.3	19	1	US-10-176-884-44	Sequence 44, Appli
C 36	17	4.2	17	1	US-08-635-820A-2	Sequence 2, Appli
C 37	17	4.2	17	1	US-09-100-104-2	Sequence 2, Appli
38	17	4.2	17	1	US-10-054-295-132	Sequence 132, App
39	17	4.2	17	1	US-09-438-486A-132	Sequence 132, App
40	17	4.2	17	1	US-09-979-275A-7	Sequence 7, Appli
C 41	16.4	4.1	18	1	US-09-018-584A-69	Sequence 69, Appli
C 42	16.4	4.1	18	1	US-09-784-423-69	Sequence 69, Appli
43	16.2	4.0	18	1	US-09-981-397A-1	Sequence 1, Appli
44	16	4.0	16	1	US-10-203-780-9	Sequence 9, Appli
C 45	16	4.0	16	1	US-10-054-295-131	Sequence 131, App
C 46	16	4.0	16	1	US-09-438-486A-131	Sequence 131, App
47	16	4.0	16	1	US-10-227-001-21	Sequence 21, Appli
48	16	4.0	16	1	US-10-072-975-9	Sequence 9, Appli
49	16	4.0	17	1	US-09-730-559B-107	Sequence 107, App
50	15	3.7	15	1	US-09-850-982B-4	Sequence 4, Appli
51	15	3.7	15	1	US-09-793-146-54	Sequence 54, Appli
52	15	3.7	15	1	US-09-793-146-55	Sequence 55, Appli
53	15	3.7	15	1	US-10-227-001-23	Sequence 23, Appli
54	15	3.7	15	1	US-10-384-450A-5	Sequence 5, Appli
55	15	3.7	15	1	US-10-072-975-10	Sequence 10, Appli
56	15	3.7	16	1	US-10-227-001-20	Sequence 20, Appli
57	14.4	3.6	16	1	US-09-479-005A-260	Sequence 260, App
58	14	3.5	15	1	US-10-227-001-24	Sequence 24, Appli

ALIGNMENTS

RESULT 1
US-08-599-252-35/c
; Sequence 35, Application US/08599252
; Patent No. 5705343
; GENERAL INFORMATION:
; APPLICANT: DRAYNA, DENNIS T.
; APPLICANT: FEDER, JOHN N.
; APPLICANT: GRIKKE, ANDREAS
; APPLICANT: KIMMEL, BRUCE E.
; APPLICANT: THOMAS, WINSTON J.
; APPLICANT: WOLFF, ROGER K.
; TITLE OF INVENTION: METHOD TO DIAGNOSE HEREDITARY
; TITLE OF INVENTION: HEMOCHROMATOSIS
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/599,252
; FILING DATE: 09-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 9053-0001.21
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```

; TOPOLOGY: linear
US-08-599-252-35

Query Match          5.1%; Score 20.4; DB 1; Length 22;
Best Local Similarity 95.5%; Pred. No. 10;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      444 GCTCTGTTGCCAGCGTGGAGT 465
Db      22 GCTCTATTGCCAGCGTGGAGT 1

RESULT 2
US-08-599-252-38/c
; Sequence 38, Application US/08599252
; Patent No. 5705343
; GENERAL INFORMATION:
; APPLICANT: DRAYNA, DENNIS T.
; APPLICANT: FEDER, JOHN N.
; APPLICANT: GNIERKE, ANDREAS
; APPLICANT: KIMMEL, BRUCE E.
; APPLICANT: THOMAS, WINSTON J.
; APPLICANT: WOLFF, ROGER K.
; TITLE OF INVENTION: METHOD TO DIAGNOSE HEREDITARY
; TITLE OF INVENTION: HEMOCHROMATOSIS
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/599,252
; FILING DATE: 09-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 9053-0001.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-599-252-38

Query Match          5.1%; Score 20.4; DB 1; Length 22;
Best Local Similarity 95.5%; Pred. No. 10;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      444 GCTCTGTTGCCAGCGTGGAGT 465
Db      22 GCTCTATTGCCAGCGTGGAGT 1

RESULT 3
US-08-291-074-2
; Sequence 2, Application US/08291074
; Patent No. 5959171
; GENERAL INFORMATION:
; APPLICANT: Hyttinen, Juha-Matti

; APPLICANT: Korhonen, Veli-Pekka
; APPLICANT: Janne, Juhani
; TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE POLYPEPTIDES IN A MAMMAL'S MILK AS
; TITLE OF INVENTION: FUSION PROTEINS THAT ARE LESS ACTIVE THAN THE FREE
; TITLE OF INVENTION: POLYPEPTIDES, OR NON-ACTIVE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Adduci, Mastriani, Schaumburg & Schill
; STREET: 1140 Connecticut Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/291,074
; FILING DATE: 17-AUG-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Kubovcik, Ronald J.
; REGISTRATION NUMBER: 25,401
; REFERENCE/DOCKET NUMBER: TUR-017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-467-6300
; TELEFAX: 202-466-2006
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-291-074-2

Query Match          5.1%; Score 20.4; DB 1; Length 22;
Best Local Similarity 95.5%; Pred. No. 10;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      464 GTGCAGTGGTGTGATCATGCT 485
Db      1 GTGCAGTGGTGTGATCATGCT 22

RESULT 4
US-08-781-891-11
; Sequence 11, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```


GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2006, 15:00:13 ; Search time 0.001 Seconds
(without alignments)
13.634 Million cell updates/sec

Title: af287270
Perfect score: 401
Sequence: 1 CAGTGGGGGATCCCATCA.....CATCCTTGGCCCTTACCCGCT 401

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 17 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 15
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1 summaries

Database : af287270_copy_6956_7356.rge4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	15.4	3.8	17	1	AX757826	ACCESSION:AX757826

ALIGNMENTS

RESULT 1
AX757826/c
LOCUS AX757826 17 bp DNA linear PAT 25-JUN-2003
DEFINITION Sequence 1147 from Patent WO03040369.
ACCESSION AX757826
VERSION AX757826.1 GI:32252442
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
REFERENCE
AUTHORS Telerman,A., Amson,R. and Tulinder,M.
TITLE Sequences involved in tumoral suppression, tumoral reversion,
apoptosis and/or viral resistance phenomena and their use as
medicines
JOURNAL Patent: WO 03040369-A 1147 15-MAY-2003;
FEATURES Molecular Engines Laboratories (FR)
source Location/Qualifiers
1..17
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Query Match 3.8%; Score 15.4; DB 1; Length 17;
Best Local Similarity 94.1%; Pred. No. 0;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 7175 CGTCCTGTGCTGAGATC 7191
||||| |||||||
Db 17 CGTCCTGGGCTGAGATC 1

Search completed: January 17, 2006, 15:00:13
Job time : 0.001 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2006, 15:01:31 ; Search time 0.001 Seconds
(without alignments)
112.280 Million cell updates/sec

Title: af287270
Perfect score: 401
Sequence: 1 CAGTGGCGGATCCCATCA.....CATCCTTGCGCCTACCGCT 401

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 0.5

Searched: 8 seqs, 140 residues

Total number of hits satisfying chosen parameters: 16

Minimum DB seq length: 15
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 8 summaries

Database : af287270_copy_6956_7356.rng4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	5.2	21	1	ADU28709 ✓ Knock-down target
C 2	21	5.2	21	1	ABE28960 Human MCOLN1 gene
C 3	20	5.0	20	1	ADB84282 Human mucolipin (M
4	16	4.0	16	1	ABE28963 Human MCOLN1 gene
C 5	15.4	3.8	17	1	ADB40824 Tumour suppression
6	14	3.5	15	1	ACD66351 Anti-HCV nucleic a
7	14	3.5	15	1	ACD66421 Anti-HCV enzymatic
8	14	3.5	15	1	ADI87738 Anti-HCV molecule

ALIGNMENTS

RESULT 1
ADU28709
ID ADU28709 standard; DNA; 21 BP.
XX AC
XX ADU28709;
XX AC
XX 27-JAN-2005 (first entry)
XX DE
XX Knock-down target sequence #2107.
XX ds; RNA production; protein production; drug development;
XX knock-down target.
XX Unidentified.
XX OS
XX WO2004094636-A1.
XX PN
XX 04-NOV-2004.
XX PD
XX 24-APR-2003; 2003WO-EP004362.

XX 24-APR-2003; 2003WO-EP004362.
XX (GALA-) GALAPAGOS GENOMICS NV.
XX PA (VSCH/) VAN DER SCHUEREN J.
XX PI Arts GJF, Lambrecht MJY, Djokic K, Clasen RJ, Mesic E;
XX PI Griffioen S, Bergs CJL;
XX DR WPI; 2004-775940/76.
XX PT New knockdown sequences, useful in lowering the amount of RNA and/or
XX PT protein production in cells used in drug development process.
XX PS Claim 11; SEQ ID NO 2123; 402pp; English.
XX CC The invention relates to a polynucleotide comprising an RNA sequence. The
XX CC polynucleotides, vector, libraries, and method are useful in lowering the
XX CC amount of RNA and/or protein production in cells used in drug development
XX CC process. The present sequence represents a knock-down target sequence.
XX SQ Sequence 21 BP; 7 A; 6 C; 5 G; 3 T; 0 U; 0 Other;

Query Match 5.2%; Score 21; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6994 ACATCCAGGAGTGAAGCACC 7014
DB 1 ACATCCAGGAGTGAAGCACC 21

RESULT 2
AEB28960/c
ID AEB28960 standard; DNA; 21 BP.
XX AC AEB28960;
XX 22-SEP-2005 (first entry)
XX DE Human MCOLN1 gene PCR primer PRI R2 SEQ ID NO:4.
XX KW mucolipidosis IV; high throughput screening; mucolipin-1; MCOLN1; PCR;
XX KW primer; ss.
XX OS Homo sapiens.
XX PN US2005153300-A1.
XX PD 14-JUL-2005.
XX PF 09-JAN-2004; 2004US-00754446.
XX PR 09-JAN-2004; 2004US-00754446.
XX PA (QUES-) QUEST DIAGNOSTICS INC.
XX PI Sun W, Hantash F;
XX DR WPI; 2005-521160/53.
XX PT Diagnosing mucolipidosis (ML) IV mutant sequence in nucleic acids, by
XX PT amplifying the nucleic acid, detecting amplified product with labeled
XX PT oligonucleotide probes via a change in fluorescence which indicates the
XX PT presence of an ML IV mutant.
XX PS Claim 1; SEQ ID NO 4; 15pp; English.
XX CC The invention relates to a method (M1) for detecting the presence of
XX CC mucolipidosis IV (ML IV) mutant sequences in a nucleic acid. (M1)
XX CC involves contacting the nucleic acid with oligonucleotide primers and
XX CC probes, conducting amplification by temperature cycling and monitoring
XX CC the accumulation of amplified nucleic acid by detecting an increase in

CC donor fluorophore fluorescence or decrease in acceptor fluorophore
 CC fluorescence, which indicates the presence of ML IV mutant sequence. Also
 CC described: (1) an oligonucleotide (ON1) comprising a sequence
 CC complementary to the coding or non-coding strand of the mucolin-1
 CC (MCOLN1) gene; and (2) a kit (K1) for amplifying sequences of ML IV
 CC sequences, comprising ON1. (M1) is useful for determining the presence of
 CC ML IV mutant sequence in a nucleic acid, and/or wild-type sequences. (M1)
 CC is useful for genetic counseling of individuals at risk for ML IV or at
 CC risk to having offspring afflicted with ML IV. The MCOLN1 gene is located
 CC on chromosome 19p13. The present sequence represents a PCR primer for the
 CC MCOLN1 gene, which is used in the exemplification of the present
 CC invention.
 SQ Sequence 21 BP; 5 A; 7 C; 5 G; 4 T; 0 U; 0 Other;

Query Match 5.2%; Score 21; DB 1; Length 21;
 Best Local Similarity 100.0%; Pred. No. 0.71;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7017 AGTGTCTTCCAGCAGCGTGAG 7037
 Db 21 AGTGTCTTCCAGCAGCGTGAG 1

RESULT 3
 ADB84282/c
 ID ADB84282 standard; DNA; 20 BP.
 XX
 AC ADB84282;
 DT 04-DEC-2003 (first entry)
 DE Human mucolin (MUCOLIPIN-1) gene AJ haplotype analysis primer #4.
 KW human; mucolin; MCOLN-1; mucopolipidosis IV; ion channel defect;
 KW immunogen; gene therapy; ss; primer; PCR; AJ haplotype.
 XX Homo sapiens.

XX US2003064363-A1.
 XX 03-APR-2003.
 XX 08-MAY-2001; 2001US-00851494.
 XX 18-AUG-2000; 2000US-0226388P.
 XX (MLFO-) ML4 FOUND & HAVARD COLLEGE.
 XX Goldin E, Slangenaupt SA, Sun M, Acierno JS;
 XX WPI; 2003-540779/51.
 XX New mucolin (MCOLN1) polypeptides and nucleic acids, useful in gene
 XX therapy for treating mucopolipidosis IV or ion channel defect, or as
 XX immunogen to generate antibodies that recognize the MCOLN1 polypeptide.
 PS Example; Page 18; 34pp; English.

XX The invention describes a nucleic acid (I) encoding a mucolin (MCOLN1)
 CC polypeptide, and sharing at least 95% sequence identity with sequence
 CC selected from 2 nucleotide sequences (designated S1 and S2, not given in
 CC the specification). A mutation of in the MCOLN1 gene encoding the MCOLN1
 CC polypeptide results in a defect in expression of a functional MCOLN1. The
 CC MCOLN1 gene is useful for the treatment of mucopolipidosis IV or
 CC ion channel defect. The MCOLN1 polypeptides are useful as immunogen to
 CC generate antibodies that recognize the MCOLN1 polypeptide. Such
 CC antibodies may be used for diagnostics and intracellular regulation of
 CC MCOLN1 activity. The polynucleotides may also be useful in gene therapy.
 CC This sequence represents a primer used to analyse the AJ haplotype of the
 CC human mucolin (MCOLN-1) gene.
 SQ Sequence 20 BP; 5 A; 7 C; 5 G; 3 T; 0 U; 0 Other;

Query Match 5.0%; Score 20; DB 1; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.94;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7018 GTGTCTTCCAGCAGCGTGAG 7037
 Db 20 GTGTCTTCCAGCAGCGTGAG 1

RESULT 4
 AEB28963
 ID AEB28963 standard; DNA; 16 BP.
 XX
 AC AEB28963;
 DT 22-SEP-2005 (first entry)

XX Human MCOLN1 gene probe DEL SEQ ID NO:7.
 DE Human MCOLN1 gene probe DEL SEQ ID NO:7.
 KW mucopolipidosis IV; high throughput screening; mucolin-1; MCOLN1; probe;
 KW ss.
 XX Homo sapiens.
 XX US2005153300-A1.
 XX 14-JUL-2005.
 XX 09-JAN-2004; 2004US-00754446.
 XX 09-JAN-2004; 2004US-00754446.
 XX (QUES-) QUEST DIAGNOSTICS INC.
 XX Sun W, Hantash F;
 XX WPI; 2005-521160/53.

XX Diagnosing mucopolipidosis (ML) IV mutant sequence in nucleic acids, by
 XX amplifying the nucleic acid, detecting amplified product with labeled
 XX oligonucleotide probes via a change in fluorescence which indicates the
 XX presence of an ML IV mutant.
 PS Claim 3; SEQ ID NO 7; 15pp; English.

XX The invention relates to a method (M1) for detecting the presence of
 CC mucopolipidosis IV (ML IV) mutant sequences in a nucleic acid. (M1)
 CC involves contacting the nucleic acid with oligonucleotide primers and
 CC probes, conducting amplification by temperature cycling and monitoring
 CC the accumulation of amplified nucleic acid by detecting an increase in
 CC donor fluorophore fluorescence or decrease in acceptor fluorophore
 CC fluorescence, which indicates the presence of ML IV mutant sequence. Also
 CC described: (1) an oligonucleotide (ON1) comprising a sequence
 CC complementary to the coding or non-coding strand of the mucolin-1
 CC (MCOLN1) gene; and (2) a kit (K1) for amplifying sequences of ML IV
 CC sequences, comprising ON1. (M1) is useful for determining the presence of
 CC ML IV mutant sequence in a nucleic acid, and/or wild-type sequences. (M1)
 CC is useful for genetic counseling of individuals at risk for ML IV or at
 CC risk to having offspring afflicted with ML IV. The MCOLN1 gene is located
 CC on chromosome 19p13. The present sequence represents a probe for the
 CC MCOLN1 gene, which is used in the exemplification of the present
 CC invention.

SQ Sequence 16 BP; 5 A; 7 C; 3 G; 1 T; 0 U; 0 Other;

Query Match 4.0%; Score 16; DB 1; Length 16;
 Best Local Similarity 100.0%; Pred. No. 2.9;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6982 AGACCCAGGCCACAT 6997
 Db 1 AGACCCAGGCCACAT 16

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: January 17, 2006, 15:03:21 ; Search time 0.001 Seconds
(without alignments)
147.568 Million cell updates/sec

Title: af287270
Perfect score: 401
Sequence: 1 CAGTGGGGGATCCCATCA.....CATCTTGGCCCTACCGCT 401

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 0.5

Searched: 10 seqs, 184 residues 20

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 15
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 10 summaries

Database : af287270_copy_6956_7356.rnpbm4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	6.2	25	1 US-11-060-756-186093	Sequence 186093,
2	25	6.2	25	1 US-11-060-756-261208	Sequence 261208,
C 3	21	5.2	21	1 US-10-754-446-4	Sequence 4, Appli
C 4	20	5.0	20	1 US-09-851-494B-11	Sequence 11, Appl
5	16	4.0	16	1 US-10-754-446-7	Sequence 7, Appli
C 6	14.4	3.6	16	1 US-09-896-324B-62	Sequence 62, Appl
C 7	14.4	3.6	16	1 US-10-236-363A-44	Sequence 44, Appl
8	14	3.5	15	1 US-09-740-332-4784	Sequence 4784, Ap
9	14	3.5	15	1 US-09-817-879-4784	Sequence 4784, Ap
10	14	3.5	15	1 US-10-669-841-7381	Sequence 7381, Ap

ALIGNMENTS

RESULT 1
US-11-060-756-186093
; Sequence 186093, Application US/11060756
; Publication No. US20050221354A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060,756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 186093
; LENGTH: 25
; TYPE: DNA
; ORGANISM: probe

US-11-060-756-186093

Query Match 6.2%; Score 25; DB 1; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6991 CCCACATCCAGGAGTGTAAAGCACCC 7015
Db 1 CCCACATCCAGGAGTGTAAAGCACCC 25

RESULT 2

US-11-060-756-261208
; Sequence 261208, Application US/11060756
; Publication No. US20050221354A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060,756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 261208
; LENGTH: 25
; TYPE: DNA
; ORGANISM: probe
US-11-060-756-261208

Query Match 6.2%; Score 25; DB 1; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6989 GGGCCACATCCAGGAGTGTAAAGCAC 7013
Db 1 GGGCCACATCCAGGAGTGTAAAGCAC 25

RESULT 3

US-10-754-446-4/c
; Sequence 4, Application US/10754446
; Publication No. US20050153300A1
; GENERAL INFORMATION:
; APPLICANT: SUN, WEIMIN
; APPLICANT: HANTASH, FERAS
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DETECTION OF
; FILE REFERENCE: 034827-2301
; CURRENT APPLICATION NUMBER: US/10/754,446
; CURRENT FILING DATE: 2004-01-09
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 4
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-10-754-446-4

Query Match 5.2%; Score 21; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7017 AGTGTCTTCCAGCAGCGGTGAG 7037
Db 21 AGTGTCTTCCAGCAGCGGTGAG 1

RESULT 4

US-09-851-494B-11/c

```
; Sequence 11, Application US/09851494B
; Publication No. US20030064363A1
; GENERAL INFORMATION:
; APPLICANT: ML4 Foundation
; APPLICANT: Goldin, Ehud
; APPLICANT: Slaugenhaupt, Susan A.
; APPLICANT: Sun, Mei
; APPLICANT: Aclerno, James S.
; TITLE OF INVENTION: A Gene Encoding A New TRP Channel is Mutated in Mucopolipidosis IV
; FILE REFERENCE: 3394/1H557US1
; CURRENT APPLICATION NUMBER: US/09/851.494B
; CURRENT FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PCR primer
US-09-851-494B-11

Query Match          5.0%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7018 GTGCTTCCAGCGGTGAG 7037
Db 20 GTGCTTCCAGCGGTGAG 1

RESULT 5
US-10-754-446-7
; Sequence 7, Application US/10754446
; Publication No. US2005015300A1
; GENERAL INFORMATION:
; APPLICANT: SUN, WEIMIN
; APPLICANT: HANTASH, FERAS
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DETECTION OF
; FILE REFERENCE: 034827-2301
; CURRENT APPLICATION NUMBER: US/10/754.446
; CURRENT FILING DATE: 2004-01-09
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 7
; LENGTH: 16
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Probe
US-10-754-446-7

Query Match          4.0%; Score 16; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6982 AGACCCAGGCCACAT 6997
Db 1 AGACCCAGGCCACAT 16

RESULT 6
US-09-896-324B-62/c
; Sequence 62, Application US/09896324B
; Publication No. US20030148276A1
; GENERAL INFORMATION:
; APPLICANT: Li, Bi-yu
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION, SEPARATION AND QUANTITATIVE MEASUREMENTS OF
; FILE REFERENCE: 45163-1008
; CURRENT APPLICATION NUMBER: US/09/896.324B
; CURRENT FILING DATE: 2002-11-04
```

```
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62
; LENGTH: 16
; TYPE: DNA
; ORGANISM: CD18-Bsaj I-ga
US-09-896-324B-62

Query Match          3.6%; Score 14.4; DB 1; Length 16;
Best Local Similarity 93.8%; Pred. No. 5.9;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7176 GTCCTGTGCTGAGATC 7191
Db 16 GTCAGTGTCTGAGATC 1

RESULT 7
US-10-236-363A-44/c
; Sequence 44, Application US/10236363A
; Publication No. US20030165923A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Xun
; APPLICANT: Li, Bi-yu
; APPLICANT: Shi, Liang
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF GENETIC MARKERS
; FILE REFERENCE: TM0011-CIP
; CURRENT APPLICATION NUMBER: US/10/236.363A
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 09/896.324
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,596
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 16
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: adapter oligonucleotide
US-10-236-363A-44

Query Match          3.6%; Score 14.4; DB 1; Length 16;
Best Local Similarity 93.8%; Pred. No. 5.9;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7176 GTCCTGTGCTGAGATC 7191
Db 16 GTCAGTGTCTGAGATC 1

RESULT 8
US-09-740-332-4784
; Sequence 4784, Application US/09740332
; Publication No. US20030125270A1
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals Inc.
; TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or Conditions Related to
; FILE REFERENCE: RPI 400/003
; CURRENT APPLICATION NUMBER: US/09/740.332
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 9704
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4784
; LENGTH: 15
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: oligonucleotide substrate
```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2006, 15:04:33 ; Search time 0.001 Seconds
(without alignments)
63.358 Million cell updates/sec

Title: af287270
Perfect score: 401
Sequence: 1 CAGTGGGGGATCCCATCA.....CATCTTGGCCCTACCGCT 401

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 0.5

Searched: 4 seqs, 79 residues

Total number of hits satisfying chosen parameters: 8

Minimum DB seq length: 15
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 4 summaries

Database : af287270_copy_6956_7356.rnpbn4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	20	5.0	21	1	US-10-310-914A-66074
					Sequence 66074, A
	2	18.4	4.6	20	1
c 3	18	4.5	20	1	US-10-310-914A-253105
					Sequence 253105, A
	4	17	4.2	18	1
					US-10-310-914A-66019
					Sequence 918753, A

ALIGNMENTS

RESULT 1
US-10-310-914A-66074/c
; Sequence 66074, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
; APPLICANT: Bentwich, Isaac
; APPLICANT: Shiler, Kvuzat
; TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
; TITLE OF INVENTION: uses thereof
; FILE REFERENCE: 06087.0200.CPUS01
; CURRENT APPLICATION NUMBER: US/10/310,914A
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 1388402
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 66074
; LENGTH: 21
; TYPE: RNA
; ORGANISM: Human
US-10-310-914A-66074

Query Match 5.0%; Score 20; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.92;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7057 CACTGACCAGGGGCCCTGGC 7076
Db 20 CACTGACCAGGGGCCCTGGC 1
|||||
|||||

RESULT 2

US-10-310-914A-253105
; Sequence 253105, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
; APPLICANT: Bentwich, Isaac
; APPLICANT: Shiler, Kvuzat
; TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
; TITLE OF INVENTION: uses thereof
; FILE REFERENCE: 06087.0200.CPUS01
; CURRENT APPLICATION NUMBER: US/10/310,914A
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 1388402
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 253105
; LENGTH: 20
; TYPE: RNA
; ORGANISM: Human
US-10-310-914A-253105

Query Match 4.6%; Score 18.4; DB 1; Length 20;
Best Local Similarity 90.0%; Pred. No. 1.5;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 7038 CCCCTGAGCCCCGACCAGC 7057
Db 1 CCCCTGAGCCCCGACCAGC 20
|||||
|||||

RESULT 3

US-10-310-914A-66019/c
; Sequence 66019, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
; APPLICANT: Bentwich, Isaac
; APPLICANT: Shiler, Kvuzat
; TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
; TITLE OF INVENTION: uses thereof
; FILE REFERENCE: 06087.0200.CPUS01
; CURRENT APPLICATION NUMBER: US/10/310,914A
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 1388402
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 66019
; LENGTH: 20
; TYPE: RNA
; ORGANISM: Human
US-10-310-914A-66019

Query Match 4.5%; Score 18; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7057 CACTGACCAGGGGCCCTG 7074
Db 18 CACTGACCAGGGGCCCTG 1
|||||
|||||

RESULT 4

US-10-310-914A-918753
; Sequence 918753, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
; APPLICANT: Bentwich, Isaac
; APPLICANT: Shiler, Kvuzat
; TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
; TITLE OF INVENTION: uses thereof
; FILE REFERENCE: 06087.0200.CPUS01

```
; CURRENT APPLICATION NUMBER: US/10/310,914A
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 1388402
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 918753
; LENGTH: 18
; TYPE: RNA
; ORGANISM: Human
US-10-310-914A-918753
```

```
Query Match      4.2%; Score 17; DB 1; Length 18;
Best Local Similarity 94.1%; Pred. No. 2.4;
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 7189 ATCCCCCAGCCCCAGA 7205
    |:|||||:|||||
Db 2 AUCCCCCAAGCCCCAGA 18
```

```
Search completed: January 17, 2006, 15:04:33
Job time : 0.001 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2006, 14:51:12 ; Search time 1 Seconds
(without alignments)
1.451 Million cell updates/sec

Title: af287270
Perfect score: 401
Sequence: 1 GGGTCTCAGCAGGACAA.....TAGCTCACTGAGCCTCGAC 401

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 0.5

Searched: 101 seqs, 1809 residues

Total number of hits satisfying chosen parameters: 202

Minimum DB seq length: 15
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 101 summaries

Database : af287270_copy_100_500.rge4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23.4	5.8	25	1	AX614112
2	20.4	5.1	22	1	AR076805
3	20.4	5.1	22	1	AR242947
4	20.4	5.1	22	1	AR345130
5	20.4	5.1	22	1	AX384999
6	20.4	5.0	20	1	BD233827
7	20.4	5.0	20	1	AR321577
8	20.4	5.0	20	1	AR542568
9	19.4	4.8	21	1	AR345126
10	19.4	4.8	21	1	AX741051
11	19.4	4.7	19	1	CS010199
12	19.4	4.7	19	1	CS028041
13	19.4	4.7	19	1	CS028508
14	19.4	4.7	19	1	AX114983
15	19.4	4.7	20	1	AR152875
16	19.4	4.7	20	1	BD134331
17	19.4	4.7	20	1	AR630295
18	19.4	4.7	20	1	AR630295
19	19.4	4.7	20	1	AR634606
20	18.4	4.6	20	1	AR162414
21	18.4	4.6	20	1	BD138315
22	18.4	4.6	20	1	BD138316
23	18.4	4.6	20	1	I31362
24	18.4	4.6	20	1	AR320228
25	18.4	4.6	20	1	AR370243
26	18.2	4.5	19	1	AR074597
27	18.2	4.5	19	1	AR083936
28	18.2	4.5	19	1	I23816
29	18.2	4.5	19	1	I29970
30	18.4	4.5	18	1	AR576394
31	18.4	4.5	19	1	CS092555
32	18.4	4.5	19	1	CS092721
33	18.4	4.5	19	1	AX133851

ALIGNMENTS

34	17.4	4.3	19	1	AR148945
35	17.4	4.3	19	1	CS027943
c 36	17.4	4.3	19	1	CS028410
c 37	17.4	4.3	19	1	AX116706
c 38	17.4	4.3	19	1	AR044034
c 39	17.4	4.3	19	1	AR592720
c 40	17.4	4.3	19	1	AR571818
41	17.4	4.3	19	1	AX761572
42	17.4	4.3	18	1	AR576395
43	17.4	4.3	18	1	AR576396
44	17.4	4.3	18	1	AR585363
45	17.4	4.3	18	1	AR612296
c 46	17.4	4.3	18	1	AR612299
c 47	17.4	4.3	18	1	AR637194
c 48	16.4	4.1	18	1	AR154019
c 49	16.4	4.1	18	1	BD130125
c 50	16.4	4.1	18	1	AR565243
51	16.4	4.1	18	1	AR575575
52	16.4	4.1	18	1	AR069644
53	16.2	4.0	18	1	AR633675
54	16.4	4.0	16	1	CS129780
55	16.4	4.0	16	1	AR584046
c 56	16.4	4.0	16	1	AR592719
57	16.4	4.0	16	1	AR642242
58	16.4	4.0	17	1	AR597132
59	16.4	4.0	17	1	AR597133
60	16.4	4.0	17	1	AR614733
c 61	16.4	4.0	17	1	AX671817
62	16.4	4.0	17	1	AX692567
63	16.4	4.0	17	1	AX692568
c 64	16.4	4.0	17	1	AX722591
c 65	16.4	4.0	17	1	AX761262
66	15.4	3.8	17	1	BD202917
67	15.4	3.8	17	1	BD202918
68	15.4	3.8	17	1	AX597131
c 69	15.4	3.8	17	1	AX671819
c 70	15.4	3.8	17	1	AX672347
c 71	15.4	3.8	17	1	AX672932
c 72	15.4	3.8	17	1	AX673646
73	15.4	3.8	17	1	AX729460
c 74	15.4	3.8	17	1	AX729877
75	15.4	3.8	17	1	AX730216
c 76	15.4	3.8	17	1	AX732723
c 77	15.4	3.8	17	1	AX732731
c 78	15.4	3.8	17	1	AX733412
c 79	15.4	3.8	17	1	AX734071
c 80	15.4	3.8	17	1	AX734153
c 81	15.4	3.8	17	1	AX734964
c 82	15.4	3.8	17	1	AX737636
c 83	15.4	3.8	17	1	AX737828
c 84	15.4	3.8	17	1	AX738556
c 85	15.4	3.8	17	1	AX739093
c 86	15.4	3.8	17	1	AX758145
c 87	15.4	3.8	17	1	AX760525
c 88	15.4	3.8	17	1	AX761520
89	15.4	3.7	15	1	CQ971639
90	15.4	3.7	15	1	CS002306
91	15.4	3.7	15	1	CS002308
92	15.4	3.7	15	1	CS002310
c 93	15.4	3.7	15	1	CS048833
c 94	15.4	3.7	15	1	CS074138
95	15.4	3.7	15	1	AR581375
96	15.4	3.7	15	1	AR612295
c 97	15.4	3.7	15	1	AR612298
98	15.4	3.7	15	1	AR630722
99	14.4	3.6	16	1	AR436001
100	14.4	3.5	15	1	CS002304
101	14.4	3.5	15	1	CS002312

RESULT 1
AX614112
LOCUS AX614112 25 bp DNA linear PAT 17-FEB-2003
DEFINITION Sequence 5137 from Patent WO02072882.
ACCESSION AX614112
VERSION AX614112.1 GI:28409541
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Cullen, P. and Seedorf, U.
TITLE Coronary chip
JOURNAL Patent: WO 02072882-A 5137 19-SEP-2002;
OGHAM GmbH (DE)
FEATURES
source Location/Qualifiers
1..25
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
Query Match 5.8%; Score 23.4; DB 1; Length 25;
Best Local Similarity 96.0%; Pred. No. 8.9;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 444 GCTCTGTGCCCCAGGCTGGAGTGCA 468
Db 1 GCTCTGTGCCCCAGGCTGGAGTGCA 25
RESULT 2
AR076805
LOCUS AR076805 22 bp DNA linear PAT 30-AUG-2000
DEFINITION Sequence 2 from patent US 5959171.
ACCESSION AR076805
VERSION AR076805.1 GI:10003551
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 22)
AUTHORS Hyttinen, J.-M., Korhonen, V.-P., and Janne, J.
TITLE Method for the production of biologically active polypeptides in a mammal's
JOURNAL Patent: US 5959171-A 2 28-SEP-1999;
FEATURES
source Location/Qualifiers
1..22
/organism="unknown"
/mol_type="unassigned DNA"
Query Match 5.1%; Score 20.4; DB 1; Length 22;
Best Local Similarity 95.5%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 464 GTGCAGTGGTGTCATCAGCT 485
Db 1 GTGCAGTGGTGTCATCAGCT 22
RESULT 3
AR242947
LOCUS AR242947 22 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 93 from patent US 6475739.
ACCESSION AR242947
VERSION AR242947.1 GI:27289609
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 22)
AUTHORS Brunkow, M.E., Proll, S., Paepers, B. and Staehling-Hampton, K.

TITLE Methods for identifying genomic deletions
JOURNAL Patent: US 6475739-A 93 05-NOV-2002;
Celltech R&D, Inc.; Bothell, WA
FEATURES
source Location/Qualifiers
1..22
/organism="unknown"
/mol_type="genomic DNA"
Query Match 5.1%; Score 20.4; DB 1; Length 22;
Best Local Similarity 95.5%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 455 CAGGCTGGAGTGCAGTGGTGTG 476
Db 1 CAGGCTGGAGTGCAGTGGTGTG 22
RESULT 4
AR345130
LOCUS AR345130 22 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 11 from patent US 6583112.
ACCESSION AR345130
VERSION AR345130.1 GI:33741766
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 22)
AUTHORS Fu, Y.-H., Yu, C.-E., Oshima, J., Mulligan, J.T. and Schellenberg, G.D.
TITLE Gene products related to Werner's syndrome
JOURNAL Patent: US 6583112-A 11 24-JUN-2003;
FEATURES
source Location/Qualifiers
1..22
/organism="unknown"
/mol_type="genomic DNA"
Query Match 5.1%; Score 20.4; DB 1; Length 22;
Best Local Similarity 95.5%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 463 AGTGCAGTGGTGTCATCAGC 484
Db 1 AGTGCAGTGGTGTCATCAGC 22
RESULT 5
AX384999
LOCUS AX384999 22 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 93 from Patent WO0210455.
ACCESSION AX384999
VERSION AX384999.1 GI:19578127
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 other sequences; artificial sequences.
AUTHORS Brunkow, M.E., Proll, S. and Paepers, B.
TITLE Methods for identifying genomic deletions
JOURNAL Patent: WO 0210455-A 93 07-FEB-2002;
Celltech R & D, Inc. (US); Straehling-Hampton, Karen (US)
FEATURES
source Location/Qualifiers
1..22
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="PCR primer"
Query Match 5.1%; Score 20.4; DB 1; Length 22;
Best Local Similarity 95.5%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 455 CAGGCTGGAGTGCAGTGGTGTG 476

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 15:49:05 ; Search time 736.68 Seconds
(without alignments)
1157.425 Million cell updates/sec

Title: US-10-754-446-6

Perfect score: 15
Sequence: 1 ctgccccgggtacct 15

Scoring table:
IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sts.*
11: gb_sy.*
12: gb_un.*
13: gb_vi.*
14: gb_htg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15	100.0	1935	6	CQ750386 Sequence
2	15	100.0	65347	14	AC130385
3	15	100.0	127168	8	HS1163J1
4	15	100.0	194763	14	AC150852
5	15	100.0	216844	14	AC130788
6	15	100.0	218893	14	AC150547
7	15	100.0	221544	14	AC127862
8	15	100.0	224638	14	AC135282
9	15	100.0	235008	14	AC103239
10	15	100.0	239553	14	AC123201
11	15	100.0	248860	14	AC115225
12	15	100.0	320746	14	AC128373
13	14	93.3	424	10	BV105141
14	14	93.3	807	6	BD162787
15	14	93.3	807	6	AX120670
16	14	93.3	930	6	AX064573
17	14	93.3	937	6	AX764219
18	14	93.3	1323	6	AR547912

19	14	93.3	1486	15	YSPADH	J01341	fisoon yeas
20	14	93.3	1985	15	AB000261	AB000261	Aspergill
21	14	93.3	3369	15	AB000262	AB000262	Aspergill
22	14	93.3	7319	6	AR227050	AR227050	Sequence
23	14	93.3	9999	15	SPCC13B11	AL032681	S.pombe c
24	14	93.3	11304	1	AE012358	AE012358	Xanthomon
25	14	93.3	12403	1	AE001706	AE001706	Thermotog
26	14	93.3	62398	14	AC166844	AC166844	Bos tauru
27	14	93.3	69517	14	AC036117	AC036117	Homo sapi
28	14	93.3	69517	14	AC036117	AC036117	Homo sapi
29	14	93.3	79925	8	AC093719	AC093719	Homo sapi
30	14	93.3	100000	8	AP000205	AP000205	Homo sapi
31	14	93.3	110000	1	AE017340_22	Continuation (23 o	
32	14	93.3	110000	1	BA000036_05	Continuation (6 of	
33	14	93.3	110000	1	CP000050_19	Continuation (20 o	
34	14	93.3	110000	14	AC095857_3	Continuation (4 of	
35	14	93.3	110000	14	CT005272_07	Continuation (8 of	
36	14	93.3	110000	15	AE017342_11	Continuation (12 o	
37	14	93.3	110000	15	AE017342_12	Continuation (13 o	
38	14	93.3	110000	15	AE017348_02	Continuation (3 of	
39	14	93.3	116409	9	AC131118	AC131118	Mus muscu
40	14	93.3	139697	14	AC134207	AC134207	Rattus no
41	14	93.3	140176	8	AC022188	AC022188	Homo sapi
42	14	93.3	140219	9	AL603714	AL603714	Mouse DNA
43	14	93.3	143509	14	AC068683	AC068683	Homo sapi
44	14	93.3	143733	14	AC152465	AC152465	Dasybus n
45	14	93.3	155628	14	AC027810	AC027810	Homo sapi

ALIGNMENTS

RESULT 1	CQ750386	Sequence 36320 from Patent WO02068579.	1935 bp	DNA	linear	PAT 03-FEB-2004
LOCUS	CQ750386					
DEFINITION	CQ750386					
ACCESSION	CQ750386.1	GI:42381467				
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
FEATURES						
source						
ORIGIN						
Query Match						
Best Local Similarity						
Matches						
Qy						
Db						
RESULT 2						
LOCUS						
DEFINITION						
ACCESSION						

```

AC130385.4 GI:23683263
KEYWORDS
SOURCE      HTGS PHASE0.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Suarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE
1 (bases 1 to 65347)
  Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
  Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
  Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
  Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
  Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
  Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
  Horton,L., Hulne,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
  Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
  Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
  McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
  Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
  O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
  Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
  Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
  Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
  Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
  Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
  Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (09-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 65347)
  Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
  Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
  Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
  Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
  Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
  Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
  Horton,L., Hulne,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
  Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
  Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
  McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
  Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
  O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
  Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
  Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
  Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
  Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
  Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
  Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (10-OCT-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Oct 10, 2002 this sequence version replaced gi:23322704.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L27996
Center clone name: 49_B_7
-----
* NOTE: This record contains 83 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for

```

```

* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 1
* 681: contig of 681 bp in length
* 682
* 781: gap of 100 bp
* 782
* 1477: contig of 696 bp in length
* 1478
* 1577: gap of 100 bp
* 1578
* 2269: contig of 692 bp in length
* 2270
* 2369: gap of 100 bp
* 2370
* 3056: contig of 687 bp in length
* 3057
* 3156: gap of 100 bp
* 3157
* 3835: contig of 679 bp in length
* 3836
* 3935: gap of 100 bp
* 3936
* 4629: contig of 694 bp in length
* 4630
* 4729: gap of 100 bp
* 4730
* 5426: contig of 697 bp in length
* 5427
* 5526: gap of 100 bp
* 5527
* 6218: contig of 692 bp in length
* 6219
* 6318: gap of 100 bp
* 6319
* 7003: contig of 685 bp in length
* 7004
* 7103: gap of 100 bp
* 7104
* 7798: contig of 695 bp in length
* 7799
* 7898: gap of 100 bp
* 7899
* 8584: contig of 686 bp in length
* 8585
* 8684: gap of 100 bp
* 8685
* 9370: contig of 686 bp in length
* 9371
* 9470: gap of 100 bp
* 9471
* 10159: contig of 689 bp in length
* 10160
* 10259: gap of 100 bp
* 10260
* 10954: contig of 695 bp in length
* 10955
* 11054: gap of 100 bp
* 11055
* 11740: contig of 686 bp in length
* 11741
* 11840: gap of 100 bp
* 11841
* 12487: contig of 647 bp in length
* 12488
* 12587: gap of 100 bp
* 12589
* 13282: contig of 695 bp in length
* 13283
* 13382: gap of 100 bp
* 13383
* 14089: contig of 687 bp in length
* 14090
* 14169: gap of 100 bp
* 14170
* 14855: contig of 686 bp in length
* 14856
* 14955: gap of 100 bp
* 14956
* 15633: contig of 678 bp in length
* 15634
* 15733: gap of 100 bp
* 15734
* 16426: contig of 693 bp in length
* 16427
* 16526: gap of 100 bp
* 16527
* 17198: contig of 672 bp in length
* 17199
* 17298: gap of 100 bp
* 17299
* 17983: contig of 685 bp in length
* 17984
* 18083: gap of 100 bp
* 18084
* 18777: contig of 694 bp in length
* 18778
* 18877: gap of 100 bp
* 18878
* 19548: contig of 671 bp in length
* 19549
* 19648: gap of 100 bp
* 19649
* 20345: contig of 697 bp in length
* 20346
* 20445: gap of 100 bp
* 20446
* 21138: contig of 693 bp in length
* 21139
* 21238: gap of 100 bp
* 21239
* 21942: contig of 704 bp in length
* 21943
* 22042: gap of 100 bp
* 22043
* 22751: contig of 709 bp in length
* 22752
* 22851: gap of 100 bp
* 22852
* 23545: contig of 694 bp in length
* 23546
* 23645: gap of 100 bp
* 23646
* 24330: contig of 685 bp in length
* 24331
* 24430: gap of 100 bp
* 24431
* 25092: contig of 662 bp in length
* 25093
* 25192: gap of 100 bp
* 25193
* 25893: contig of 701 bp in length
* 25894
* 25993: gap of 100 bp
* 25994
* 26679: contig of 686 bp in length

```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 15:55:50 ; Search time 1677.48 Seconds
(without alignments)
418.369 Million cell updates/sec

Title: US-10-754-446-6

Perfect score: 15

Sequence: 1 ctgccacggtacct 15

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_ges1:*
10: gb_ges2:*
11: gb_ges3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15	100.0	332	2	BF353711 QV2-HT069
2	15	100.0	398	6	CB707082 AMGNNUC:N
3	15	100.0	615	6	CB579676 AMGNNUC:N
4	15	100.0	674	10	CB897726 pastbac05
5	15	100.0	680	7	CK845557 UI-R-BJ1
6	15	100.0	696	8	DN105468 1101440 M
7	15	100.0	697	8	CV796826 UI-R-EB1
8	15	100.0	748	7	CO565750 AGENCOURT
9	15	100.0	763	8	CV991014 IPCGPr1.5
10	15	100.0	905	7	CN248276 EST014180
11	14	93.3	198	7	CO261528 4133787 B
12	14	93.3	303	2	BF394984 UI-R-CM0
13	14	93.3	303	9	AQ918623 RPCI-23-2
14	14	93.3	318	8	DN278032 1157316 M
15	14	93.3	346	1	AW710109 eIC09ne.f
16	14	93.3	373	5	BU994782 HM08C24r
17	14	93.3	405	2	BI187267 alG04fs.r
18	14	93.3	445	1	AU007564 AU007564
19	14	93.3	451	6	CF192358 14d08j2.f
20	14	93.3	458	3	BI679825 457072 MA
21	14	93.3	461	1	AW358990 44588 MAR
22	14	93.3	471	1	AW342926 fJ80910.Y

C	23	14	93.3	477	6	CF708620	CCAN62TF
	24	14	93.3	479	1	AU008740	AU008740
	25	14	93.3	492	3	BI538165	BI538165
	26	14	93.3	514	3	BI899044	BI899044
	27	14	93.3	515	6	CF364476	CF364476
	28	14	93.3	519	2	BE235591	BE235591
	29	14	93.3	526	10	CE455507	CE455507
	30	14	93.3	528	1	AV434883	AV434883
	31	14	93.3	529	2	BE013702	BE013702
	32	14	93.3	547	3	BI343534	BI343534
	33	14	93.3	548	1	AV439442	AV439442
	34	14	93.3	549	2	BF198928	BF198928
	35	14	93.3	551	9	AZ840147	AZ840147
	36	14	93.3	552	6	CF673281	CF673281
	37	14	93.3	554	1	AV665471	AV665471
	38	14	93.3	578	3	BI774227	BI774227
	39	14	93.3	583	5	BX670522	BX670522
	40	14	93.3	599	6	CB459576	CB459576
	41	14	93.3	621	8	CX804839	CX804839
	42	14	93.3	630	6	CB461236	CB461236
	43	14	93.3	633	10	CW101525	CW101525
	44	14	93.3	637	8	DN128186	DN128186
	45	14	93.3	648	8	DN746616	DN746616

ALIGNMENTS

RESULT 1
BF353711
LOCUS QV2-HT0698-110700-267-a01 HT0698 Homo sapiens cdna, mRNA linear EST 22-NOV-2000
DEFINITION BF353711
ACCESSION BF353711
VERSION BF353711.1 GI:11312785
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 332)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Ngai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Frudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=QV2&t2=QV2-HT0698-
110700-267-a01&t3=2000-07-11&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 11
High quality sequence stop: 331.
Location/Qualifiers
1. 332
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0698"

/notes="Organ: head neck; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 100.0%; Score 15; DB 2; Length 332;
Best Local Similarity 100.0%; Pred. NO. 1.7e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCCACGGTACTT 15
|||||
Db 236 CTGCCACGGTACTT 250

RESULT 2

CB707082 398 bp mRNA linear EST 10-APR-2003
LOCUS AMGNNUC.NRHV5-00217-H7-A W Rat hypothalamus (10471) Rattus
DEFINITION norvegicus cDNA clone nrh5-00217-h7 5', mRNA sequence.
ACCESSION CB707082
VERSION
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM

REFERENCE

1 (bases 1 to 398)
AUTHORS Angen EST Program
TITLES Angen Rat EST Program
JOURNAL Unpublished (2003)
COMMENT Contact: Dan Fitzpatrick
Angen, Inc
One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00217 row: h column: 7.

FEATURES

source
Location/Qualifiers
1. .398
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="nrh5-00217-h7"
/notes="Vector: pSPORT1; Site 1: SmaI; Site 2: NotI; W Rat hypothalamus adult female Wistar rat avg. insert size 2.3 kb fraction 6 and 7"

ORIGIN

Query Match 100.0%; Score 15; DB 6; Length 398;
Best Local Similarity 100.0%; Pred. NO. 1.7e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCCACGGTACTT 15
|||||
Db 40 CTGCCACGGTACTT 54

RESULT 3

CB579676 615 bp mRNA linear EST 03-APR-2003
LOCUS AMGNNUC.NRHV5-00106-F3-A W Rat hypothalamus (10471) Rattus
DEFINITION norvegicus cDNA clone nrh5-00106-f3 5', mRNA sequence.
ACCESSION CB579676
VERSION
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM

REFERENCE

1 (bases 1 to 615)
AUTHORS Angen EST Program
TITLES Angen Rat EST Program
JOURNAL Unpublished (2003)
COMMENT Contact: Dan Fitzpatrick
Angen, Inc
One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00106 row: f column: 3.

Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
1 (bases 1 to 615)
Angen EST Program
Angen Rat EST Program
Unpublished (2003)
Contact: Dan Fitzpatrick
Angen, Inc
One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00106 row: f column: 3.

REFERENCE
AUTHORS
TITLES
JOURNAL
COMMENT

FEATURES
source
Location/Qualifiers
1. .615
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="nrh5-00106-f3"
/notes="Vector: pSPORT1; Site 1: SmaI; Site 2: NotI; W Rat hypothalamus adult female Wistar rat avg. insert size 2.3 kb fraction 6 and 7"

ORIGIN

Query Match 100.0%; Score 15; DB 6; Length 615;
Best Local Similarity 100.0%; Pred. NO. 1.8e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCCACGGTACTT 15
|||||
Db 145 CTGCCACGGTACTT 159

RESULT 4

CG897726 674 bp DNA linear GSS 08-DEC-2003
LOCUS pastbac058xx02.bl.ab1 Res147 1 Pasteuria penetrans genomic 5',
DEFINITION genomic survey sequence.
ACCESSION CG897726
VERSION
KEYWORDS
SOURCE Pasteuria penetrans
ORGANISM

REFERENCE
AUTHORS
TITLES
JOURNAL
COMMENT

1 (bases 1 to 674)
Opperman, C.H., Davies, K.G., Sosinski, B.R., Waterman, J. and Burke, M.
Unpublished Data
Unpublished (2003)
Contact: Opperman CH
Center for the Biology of Nematode Parasitism and Nematode Interactions Unit
North Carolina State University and Rothamsted Research, Ltd.
Box 7253, NCSU, Raleigh, NC 27606, USA
Email: warthog@unity.ncsu.edu
Homology: e-val = 1e-51. Description = DNA ligase
(polydeoxyribonucleotide synthase [NAD+])
gi|3688229|emb|CAA09732.1| DNA ligase [Geobacillus
stearothermophilus] Homology: e-val = 1e-50. Description = DNA
ligase (polydeoxyribonucleotide synthase [NAD+]) [Bacillus
halodurans] gi|15213980|sp|Q9KF37|DNLJ_BACHD DNA ligase
(polydeoxyribonucleotide synthase [NAD+]) gi|25293998|pir|A83731
DNA ligase (polydeoxyribonucleotide synthase [NAD+]) lig [imported]
- row: k column: 2
Class: shotgun
High quality sequence stop: 674.

FEATURES
source
Location/Qualifiers
1. .674
/organism="Pasteuria penetrans"
/mol_type="genomic DNA"
/strain="Res147"
/db_xref="taxon:86005"
/clone_lib="Res147 1"

ORIGIN

Query Match 100.0%; Score 15; DB 10; Length 674;

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 14:15:14 ; Search time 203.52 Seconds
(without alignments)
491.207 Million cell updates/sec

Title: US-10-754-446-6

Perfect score: 15

Sequence: 1 ctgccacggtacct 15

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21.*
1: geneseqn1980s.*
2: geneseqn1990s.*
3: geneseqn2000s.*
4: geneseqn2001as.*
5: geneseqn2001bs.*
6: geneseqn2002as.*
7: geneseqn2002bs.*
8: geneseqn2003as.*
9: geneseqn2003bs.*
10: geneseqn2003cs.*
11: geneseqn2003ds.*
12: geneseqn2004as.*
13: geneseqn2004bs.*
14: geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15	100.0	15	14 AEB28962	Aeb28962 Human MCO
2	14	93.3	560	13 ADQ49258	Adq49258 Novel can
C 3	14	93.3	807	5 AAH65551	Aah65551 C glutami
C 4	14	93.3	930	4 AAF72180	Aaf72180 Corynebac
C 5	14	93.3	937	10 ADD13464	Add13464 C. glutam
6	14	93.3	1050	13 ADT47984	Adt47984 Bacterial
7	14	93.3	1365	10 ADC81562	Adc81562 Recombina
8	14	93.3	1422	13 ADS48605	Ads48605 Bacterial
C 9	14	93.3	1524	13 ADS56578	Ads56578 Bacterial
10	14	93.3	34980	5 AAH68525	Aah68525 C glutami
11	13.4	89.3	16	14 AEB28961	Aeb28961 Human MCO
C 12	13.4	89.3	118	5 ABV43024	Abv43024 Human pro
C 13	13.4	89.3	118	5 ABV34163	Abv34163 Human pro
14	13.4	89.3	146	12 ACH93483	Ach93483 Human gen
C 15	13.4	89.3	252	12 ADP61947	Adp61947 Maize car
C 16	13.4	89.3	265	12 ADP61965	Adp61965 Maize car
C 17	13.4	89.3	267	12 ADP61934	Adp61934 Maize car
C 18	13.4	89.3	270	12 ADP61924	Adp61924 Maize car
C 19	13.4	89.3	272	12 ADP61945	Adp61945 Maize car

C 20	13.4	89.3	273	2 AAV09248	Aav09248 Human cyt
C 21	13.4	89.3	273	2 AAV12206	Aav12206 Human ret
C 22	13.4	89.3	273	6 AAD24507	Aad24507 Human P45
C 23	13.4	89.3	273	13 ADU66854	Adu66854 hp450RAI
C 24	13.4	89.3	273	14 ADV90794	Adv90794 Human ret
C 25	13.4	89.3	273	14 ADV62611	Adv62611 Human P45
C 26	13.4	89.3	274	2 AAV09249	Aav09249 Murine cy
C 27	13.4	89.3	274	2 AAV12207	Aav12207 Mouse ret
C 28	13.4	89.3	274	6 AAD24508	Aad24508 Mouse P45
C 29	13.4	89.3	274	13 ADU66855	Adu66855 mp450RAI
C 30	13.4	89.3	274	14 ADV90795	Adv90795 Murine re
C 31	13.4	89.3	274	14 ADV62612	Adv62612 Mouse P45
C 32	13.4	89.3	278	12 ADP61931	Adp61931 Maize car
C 33	13.4	89.3	285	12 ADP61916	Adp61916 Maize car
C 34	13.4	89.3	296	12 ADP61894	Adp61894 Maize car
C 35	13.4	89.3	298	4 AAL11859	Aal11859 Human bre
C 36	13.4	89.3	304	12 ADP61910	Adp61910 Maize car
C 37	13.4	89.3	311	12 ADP61878	Adp61878 Maize car
C 38	13.4	89.3	312	12 ADP61883	Adp61883 Maize car
C 39	13.4	89.3	314	12 ADP61875	Adp61875 Maize car
C 40	13.4	89.3	315	12 ADP61873	Adp61873 Maize car
C 41	13.4	89.3	321	12 ADP61869	Adp61869 Maize car
C 42	13.4	89.3	324	12 ADP61867	Adp61867 Maize car
C 43	13.4	89.3	341	4 AAS35653	Aas35653 Human car
C 44	13.4	89.3	341	10 ADE45732	Ade45732 Human car
C 45	13.4	89.3	341	13 ADJ07150	Adj07150 Human car

ALIGNMENTS

RESULT 1

AEB28962

ID AEB28962 standard; DNA; 15 BP.

AC AEB28962;

DT 22-SEP-2005 (first entry)

DE Human MCOLN1 gene probe IVS MUT SEQ ID NO:6.

KW mucolipidosis IV; high throughput screening; mucolipin-1; MCOLN1; probe;
KW ss.

OS Homo sapiens.

PN US2005153300-A1.

PD 14-JUL-2005.

PF 09-JAN-2004; 2004US-00754446.

PR 09-JAN-2004; 2004US-00754446.

PA (QUES-) QUEST DIAGNOSTICS INC.

PI Sun W, Kantash F;

XX WPI; 2005-521160/53.

XX Diagnosing mucolipidosis (ML) IV mutant sequence in nucleic acids, by
amplifying the nucleic acid, detecting amplified product with labeled
oligonucleotide probes via a change in fluorescence which indicates the
presence of an ML IV mutant.

PS Claim 3; SEQ ID NO 6; 15pp; English.

XX The invention relates to a method (M1) for detecting the presence of
mucolipidosis IV (ML IV) mutant sequences in a nucleic acid. (M1)
involves contacting the nucleic acid with oligonucleotide primers and
probes, conducting amplification by temperature cycling and monitoring
the accumulation of amplified nucleic acid by detecting an increase in
donor fluorophore fluorescence or decrease in acceptor fluorophore

CC fluorescence, which indicates the presence of ML IV mutant sequence. Also
 CC described: (1) an oligonucleotide (ON1) comprising a sequence
 CC complementary to the coding or non-coding strand of the mucopolin-1
 CC (MCOLN1) gene; and (2) a kit (K1) for amplifying sequences of ML IV
 CC sequences, comprising ON1. (M1) is useful for determining the presence of
 CC ML IV mutant sequence in a nucleic acid, and/or wild-type sequences. (M1)
 CC is useful for genetic counseling of individuals at risk for ML IV or at
 CC risk to having offspring afflicted with ML IV. The MCOLN1 gene is located
 CC on chromosome 19p13. The present sequence represents a probe for the
 CC MCOLN1 gene, which is used in the exemplification of the present
 CC invention.
 XX
 SQ Sequence 15 BP; 2 A; 7 C; 3 G; 3 T; 0 U; 0 Other;
 Query Match 100.0%; Score 15; DB 14; Length 15;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02; Mismatches 0; Indels 0; Gaps 0;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CTGCCACGGTACCT 15
 ID ADQ49258 standard; DNA; 560 BP.
 AC ADQ49258;
 XX
 XX 21-OCT-2004 (first entry)
 XX Novel canine microarray-related DNA sequence SeqID560.
 XX canine microarray; drug screening; toxicity assay;
 XX environmental pollutant; cellular response; gene expression profile;
 KW toxic response; liver necrosis; fatty liver disease;
 KW protein adduct formation; hepatitis; dog; ds.
 XX
 OS Canis familiaris.
 XX
 XX WO2004063324-A2.
 PN 29-JUL-2004.
 XX
 XX 05-MAY-2003; 2003WO-US013853.
 XX
 XX 03-MAY-2002; 2002US-0377240P.
 PR
 XX (GENE-) GENE LOGIC INC.
 PA (PFIZ) PFIZER PROD INC.
 XX
 XX Diggins JC, Porter M, Wei T;
 XX WPI; 2004-561890/54.
 XX
 XX New isolated nucleic acid molecule, useful for drug screening and
 PT toxicity assays or for assessing the impact, including toxicity, of a
 PT compound, pharmaceutical agent or environmental pollutant on a cell or
 PT living organism.
 XX
 XX Claim 1; SEQ ID NO 560; 41pp; English.
 XX
 CC This invention is related to a novel isolated canine nucleic acid
 CC sequences and the construction of canine microarrays containing a
 CC significant portion of the canine genome. The isolated canine nucleic
 CC acid sequences of the invention may be useful for drug screening and
 CC toxicity assays. The invention is therefore useful for assessing the
 CC impact, including toxicity, of a compound, pharmaceutical agent or
 CC environmental pollutant on a cell or living organism. The methods are
 CC useful for detecting genes that are up- or down-regulated in canines in a
 CC disease state. The sequences are useful as diagnostic agents or markers
 CC to detect a cellular response in a sample individually or as part of a
 CC gene expression profile. It is also useful as a target for agents that

CC modulate gene expression or activity. The database is useful for
 CC producing electronic Northern blots that allow the user to determine the cell
 CC type or tissue in which a given gene is expressed and to allow
 CC determination of the abundance or expression level of a given gene in a
 CC particular tissue or cell. The methods are useful for determining the
 CC similarity of a toxic response to one or more individual compounds. The
 CC methods are useful for predicting at least one toxic response or the
 CC likelihood that a compound or test agent will induce various specific
 CC pathologies such as those of the liver (liver necrosis, fatty liver
 CC disease, protein adduct formation or hepatitis), those of the kidney,
 CC heart, brain or testes, or other pathologies associated with at least one
 CC of the toxins. The methods are also useful for predicting or elucidating
 CC the potential cellular pathways influenced, induced or modulated by the
 CC compound or test agent due to the similarity of the expression profile
 CC compared to the profile induced by a known toxin. The present sequence is
 CC that of a canine DNA sequence which was claimed for use during the
 CC production of a canine microarray of the invention.
 XX
 SQ Sequence 560 BP; 133 A; 168 C; 152 G; 102 T; 0 U; 5 Other;
 Query Match 93.3%; Score 14; DB 13; Length 560;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03; Mismatches 0; Indels 0; Gaps 0;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CTGCCACGGTACC 14
 ID AAH65551/c
 XX AAH65551 standard; DNA; 807 BP.
 AC AAH65551;
 XX 26-SBP-2001 (first entry)
 DT
 XX C glutamicum coding sequence fragment SEQ ID NO: 586.
 DE
 XX Corynebacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis; ds.
 XX
 OS Corynebacterium glutamicum.
 XX
 XX EP1108790-A2.
 PN
 XX 20-JUN-2001.
 PD
 XX 18-DEC-2000; 2000EP-00127688.
 XX
 XX 16-DEC-1999; 99JP-00377484.
 PR
 XX 07-APR-2000; 2000JP-00159162.
 PR
 XX 03-AUG-2000; 2000JP-00280988.
 PR
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 PA
 XX Nakagawa S, Mizoguchi H, Ando S, Havaashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 PI
 XX WPI; 2001-376931/40.
 DR
 XX P-PSDB; AAG90332.
 DR
 XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Claim 8; SEQ ID NO 586; 246pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analysing
 CC the expression profile or expression pattern of a gene derived from

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 16:41:33 ; Search time 362.28 Seconds
(without alignments)
342.389 Million cell updates/sec

Title: US-10-754-446-6

Perfect score: 15

Sequence: 1 CTGCCACGGTACCT 15

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main:

- 1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq.*
- 3: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq.*
- 4: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq.*
- 6: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq.*
- 8: /cgn2_6/ptodata/1/pubna/US10D_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubna/US10E_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15	100.0	15	9	US-10-754-446-6
2	14	93.3	25	8	US-10-719-900-215409
3	14	93.3	491	7	US-10-424-599-52366
4	14	93.3	500	4	US-09-925-065A-757224
5	14	93.3	807	3	US-09-738-626-586
6	14	93.3	937	8	US-10-494-675-31
7	14	93.3	1050	6	US-10-369-493-46422
8	14	93.3	1365	6	US-10-372-473-8
9	14	93.3	1422	6	US-10-369-493-27035
10	14	93.3	1524	6	US-10-369-493-32252
11	14	93.3	7319	5	US-10-194-163-510
12	14	93.3	3309400	3	US-09-738-626-1
13	13.4	89.3	16	9	US-10-754-446-5
14	13.4	89.3	25	10	US-11-036-317-515390
15	13.4	89.3	118	8	US-10-357-930-34181
16	13.4	89.3	146	8	US-10-357-930-43043
17	13.4	89.3	201	6	US-10-023-386-26678
18	13.4	89.3	201	8	US-10-719-993-9298
19	13.4	89.3	201	8	US-10-741-600-14149
20	13.4	89.3	201	8	US-10-741-600-14194
21	13.4	89.3	201	8	US-10-741-600-14264
22	13.4	89.3	201	8	US-10-741-600-14300
23	13.4	89.3	201	8	US-10-741-600-14338

C 24	13.4	89.3	201	8	US-10-741-600-14361	Sequence 14361, A
C 25	13.4	89.3	201	8	US-10-741-600-26703	Sequence 26703, A
C 26	13.4	89.3	201	8	US-10-741-600-49737	Sequence 49737, A
C 27	13.4	89.3	252	3	US-09-987-899-5016	Sequence 5016, Ap
C 28	13.4	89.3	265	3	US-09-987-899-5034	Sequence 5034, Ap
C 29	13.4	89.3	267	3	US-09-987-899-5003	Sequence 5003, Ap
C 30	13.4	89.3	270	3	US-09-987-899-4993	Sequence 4993, Ap
C 31	13.4	89.3	272	3	US-09-987-899-5014	Sequence 5014, Ap
C 32	13.4	89.3	273	8	US-10-855-595-33	Sequence 33, Appl
C 33	13.4	89.3	273	8	US-10-855-532-33	Sequence 33, Appl
C 34	13.4	89.3	274	8	US-10-855-595-34	Sequence 34, Appl
C 35	13.4	89.3	274	8	US-10-855-532-34	Sequence 34, Appl
C 36	13.4	89.3	278	3	US-09-987-899-5000	Sequence 5000, Ap
C 37	13.4	89.3	285	3	US-09-987-899-4985	Sequence 4985, Ap
C 38	13.4	89.3	296	3	US-09-987-899-4963	Sequence 4963, Ap
C 39	13.4	89.3	304	3	US-09-987-899-4979	Sequence 4979, Ap
C 40	13.4	89.3	310	7	US-10-242-535A-32082	Sequence 32082, A
C 41	13.4	89.3	310	7	US-10-085-783A-32082	Sequence 32082, A
C 42	13.4	89.3	311	3	US-09-987-899-4947	Sequence 4947, Ap
C 43	13.4	89.3	312	3	US-09-796-692-3501	Sequence 3501, Ap
C 44	13.4	89.3	312	3	US-09-987-899-4952	Sequence 4952, Ap
C 45	13.4	89.3	312	5	US-10-040-862-3501	Sequence 3501, Ap

ALIGNMENTS

RESULT 1
US-10-754-446-6
; Sequence 6, Application US/10754446
; Publication No. US20050153300A1
; GENERAL INFORMATION:
; APPLICANT: HANTASH, PERAS
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DETECTION OF
; TITLE OF INVENTION: MUCOLIPIDOSIS IV MUTATIONS
; FILE REFERENCE: 034827-2301
; CURRENT APPLICATION NUMBER: US/10/754,446
; CURRENT FILING DATE: 2004-01-09
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 3.2
; SEQ ID NO 6
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Probe
US-10-754-446-6

Query Match 100.0%; Score 15; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCCACGGTACCT 15
| | | | | | | | | | | | | | | |
Db 1 CTGCCACGGTACCT 15

RESULT 2
US-10-719-900-215409
; Sequence 215409, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002-11-20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 215409

```
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-215409

Query Match          93.3%; Score 14; DB 8; Length 25;
Best Local Similarity 100.0%; Pred. No. 9e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  2 TGCCACGGTACCT 15
    |||||
Db   3 TGCCACGGTACCT 16

RESULT 3
US-10-424-599-52366/c
; Sequence 52366, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 52366
; LENGTH: 491
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_182C.1
US-10-424-599-52366

Query Match          93.3%; Score 14; DB 7; Length 491;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  2 TGCCACGGTACCT 15
    |||||
Db   190 TGCCACGGTACCT 177

RESULT 4
US-09-925-065A-757224/c
; Sequence 757224, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 757224
; LENGTH: 500
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
US-09-925-065A-757224

Query Match          93.3%; Score 14; DB 4; Length 500;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  1 CTGCCACGGTACC 14
    |||||
Db   119 CTGCCACGGTACC 106

RESULT 5
US-09-738-626-586/c
; Sequence 586, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 586
; LENGTH: 807
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-586

Query Match          93.3%; Score 14; DB 3; Length 807;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  2 TGCCACGGTACCT 15
    |||||
Db   145 TGCCACGGTACCT 132

RESULT 6
US-10-494-675-31/c
; Sequence 31, Application US/10494675
; Publication No. US20050019877A1
; GENERAL INFORMATION:
; APPLICANT: Zelder, Oskar
; APPLICANT: Pompejus, Markus
; APPLICANT: Schroder, Hartwig
; APPLICANT: Kroger, Burkhard
; APPLICANT: Klopprogge, Corinna
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: Genes coding for metabolic pathway proteins
; FILE REFERENCE: BGI-163US
; CURRENT APPLICATION NUMBER: US/10/494,675
; CURRENT FILING DATE: 2004-05-04
; PRIOR APPLICATION NUMBER: PCT/EP02/12141
; PRIOR FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: DE 101 54 292.1
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 164
```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 16:59:57 ; Search time 184.56 Seconds
(without alignments)
65.779 Million cell updates/sec

Title: US-10-754-446-6

Perfect score: 15

Sequence: 1 ctgcccacggtacct 15

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 6038814 seqs, 404674181 residues

Total number of hits satisfying chosen parameters: 12077628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA New:

- 1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2.*
- 9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3.*
- 10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	14	93.3	930	7	US-11-085-822-855
C 2	13.4	89.3	201	6	US-10-995-561-41220
C 3	13.4	89.3	708	6	US-10-750-185-33670
C 4	13.4	89.3	708	6	US-10-750-623-33670
C 5	13.4	89.3	847	6	US-10-750-185-61862
C 6	13.4	89.3	847	6	US-10-750-623-61862
C 7	13.4	89.3	1372	6	US-10-750-185-57181
C 8	13.4	89.3	1372	6	US-10-750-623-57181
C 9	13.4	89.3	1404	6	US-10-517-939-303
C 10	13.4	89.3	9471	6	US-10-821-234-410
C 11	13.4	89.3	168516	7	US-11-121-086-3
C 12	13.4	89.3	1125000	6	US-10-995-561-13286
C 13	13	86.7	93	6	US-10-310-914A-17319
C 14	13	86.7	441	6	US-10-835-615-111
C 15	13	86.7	3336	6	US-10-750-185-55139
C 16	13	86.7	3336	6	US-10-750-623-55139
C 17	13	86.7	167116	7	US-11-121-086-44
C 18	13	86.7	182303	7	US-11-121-086-45
C 19	12.4	82.7	19	8	US-11-101-244-511715
C 20	12.4	82.7	19	8	US-11-101-244-1350721
C 21	12.4	82.7	19	9	US-11-083-784-511715
C 22	12.4	82.7	19	9	US-11-083-784-1350721
C 23	12.4	82.7	20	6	US-10-310-914A-1021148

C 24	12.4	82.7	22	6	US-10-310-914A-1387393
C 25	12.4	82.7	25	6	US-10-310-914A-855540
C 26	12.4	82.7	25	7	US-11-121-849-221946
C 27	12.4	82.7	25	7	US-11-121-849-564915
C 28	12.4	82.7	201	6	US-10-995-561-5198
C 29	12.4	82.7	201	6	US-10-995-561-5207
C 30	12.4	82.7	201	6	US-10-995-561-5283
C 31	12.4	82.7	201	6	US-10-995-561-5292
C 32	12.4	82.7	201	6	US-10-995-561-5362
C 33	12.4	82.7	201	6	US-10-995-561-5371
C 34	12.4	82.7	201	6	US-10-995-561-16810
C 35	12.4	82.7	201	6	US-10-995-561-16812
C 36	12.4	82.7	201	6	US-10-995-561-24233
C 37	12.4	82.7	201	6	US-10-995-561-24481
C 38	12.4	82.7	201	6	US-10-995-561-28948
C 39	12.4	82.7	201	6	US-10-995-561-29048
C 40	12.4	82.7	201	6	US-10-995-561-29049
C 41	12.4	82.7	201	6	US-10-995-561-29168
C 42	12.4	82.7	600	7	US-11-128-061-4350
C 43	12.4	82.7	625	6	US-10-750-185-54379
C 44	12.4	82.7	625	6	US-10-750-623-54379
C 45	12.4	82.7	1059	6	US-10-750-185-44993

ALIGNMENTS

RESULT 1

US-11-055-822-855/c
; Sequence 855, Application US/11055822
; Publication No. US20050260707A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Krogger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; FILE REFERENCE: BGI-121CPCN
; CURRENT APPLICATION NUMBER: US/11/055,822
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: 03/606,740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141,031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142,101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148,613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187,970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19930476.9
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931415.2
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931418.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; Remaining Prior Application data removed - See File wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1158
; SEQ ID NO 855
; LENGTH: 930
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(907)
; OTHER INFORMATION: RXN00708
US-11-055-822-855

```
Query Match      93.3%; Score 14; DB 7; Length 930;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 TGCCACGGTACCT 15
DB      245 TGCCACGGTACCT 232
|||||

RESULT 2
US-10-995-561-41220/c
; Sequence 41220, Application US/10995561
; Publication No. US2005027054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41220
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-41220

Query Match      89.3%; Score 13.4; DB 6; Length 201;
Best Local Similarity 93.3%; Pred. No. 2.1e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 CTGCCACGGTACCT 15
DB      42 CTGCCACGGGACCT 28
|||||

RESULT 3
US-10-750-185-33670/c
; Sequence 33670, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2003-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 33670
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Bovine 19866880906011
US-10-750-185-33670

Query Match      89.3%; Score 13.4; DB 6; Length 708;
Best Local Similarity 93.3%; Pred. No. 2e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 CTGCCACGGTACCT 15
DB      106 CTGCCACGGGACCT 92
|||||

RESULT 4
US-10-750-623-33670/c
; Sequence 33670, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2003-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 33670
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Bovine 19866880906011
US-10-750-623-33670

Query Match      89.3%; Score 13.4; DB 6; Length 708;
Best Local Similarity 93.3%; Pred. No. 2e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 CTGCCACGGTACCT 15
DB      106 CTGCCACGGGACCT 92
|||||

RESULT 5
US-10-750-185-61862/c
; Sequence 61862, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2003-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 61862
; LENGTH: 847
; TYPE: DNA
; ORGANISM: Bovine 19866881794277
US-10-750-185-61862

Query Match      89.3%; Score 13.4; DB 6; Length 847;
Best Local Similarity 93.3%; Pred. No. 2e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 CTGCCACGGTACCT 15
DB      177 CTGCCACGGGACCT 163
|||||

RESULT 6
US-10-750-623-61862/c
```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 16:32:47 ; Search time 54.6 Seconds
(without alignments)
488.341 Million cell updates/sec

Title: US-10-754-446-6

Perfect score: 15

Sequence: 1 ctgccccgggtacct 15

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/PC/TUS COMB.seq.*
- 7: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
- 8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
- 9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14	93.3	1323	3	US-09-248-796A-3043
2	14	93.3	7319	3	US-09-221-017B-510
3	13.4	89.3	63	2	US-08-707-237A-5
4	13.4	89.3	273	3	US-08-882-164D-33
5	13.4	89.3	273	3	US-09-668-482-33
6	13.4	89.3	274	3	US-08-882-164D-34
7	13.4	89.3	274	3	US-09-668-482-34
8	13.4	89.3	482	3	US-10-131-827-8312
9	13.4	89.3	601	3	US-09-949-016-78826
10	13.4	89.3	601	3	US-09-949-016-169541
11	13.4	89.3	601	3	US-09-949-016-169542
12	13.4	89.3	601	3	US-09-949-016-169543
13	13.4	89.3	1047	3	US-09-107-433-1277
14	13.4	89.3	1125	2	US-08-583-562B-9
15	13.4	89.3	1125	2	US-08-779-113-9
16	13.4	89.3	1143	3	US-09-583-110-2113
17	13.4	89.3	1494	2	US-08-583-562B-11
18	13.4	89.3	1494	2	US-08-779-113-11
19	13.4	89.3	1533	3	US-09-075-454-11
20	13.4	89.3	1660	3	US-09-347-650-11
21	13.4	89.3	2574	2	US-08-583-562B-1
22	13.4	89.3	2574	2	US-08-779-113-1
23	13.4	89.3	2677	3	US-08-882-164D-36
24	13.4	89.3	2677	3	US-09-668-482-36

C 25	13.4	89.3	3252	3	US-09-774-528-104	Sequence 104, App
C 26	13.4	89.3	3252	3	US-10-120-988-104	Sequence 104, App
C 27	13.4	89.3	3561	2	US-08-097-997A-12	Sequence 12, Appl
C 28	13.4	89.3	3561	3	US-08-665-574C-12	Sequence 12, Appl
C 29	13.4	89.3	3561	3	US-08-946-994-12	Sequence 12, Appl
C 30	13.4	89.3	4080	3	US-09-016-434-1353	Sequence 1353, Ap
C 31	13.4	89.3	4176	3	US-09-972-800A-17	Sequence 17, Appl
C 32	13.4	89.3	4176	3	US-09-023-655-1378	Sequence 1378, Ap
C 33	13.4	89.3	5238	3	US-09-620-312D-351	Sequence 351, App
C 34	13.4	89.3	7559	2	US-08-250-848-2	Sequence 2, Appl
C 35	13.4	89.3	11915	3	US-08-961-527-96	Sequence 96, Appl
C 36	13.4	89.3	12681	3	US-09-949-016-16576	Sequence 16576, A
C 37	13.4	89.3	15353	3	US-09-949-016-13454	Sequence 13454, A
C 38	13.4	89.3	17478	3	US-09-949-016-12194	Sequence 12194, A
C 39	13.4	89.3	17479	3	US-09-949-016-13057	Sequence 13057, A
C 40	13.4	89.3	44821	3	US-09-949-016-13764	Sequence 13764, A
C 41	13.4	89.3	67745	3	US-09-949-016-17251	Sequence 17251, A
C 42	13.4	89.3	110585	3	US-09-949-016-13427	Sequence 13427, A
C 43	13.4	89.3	160759	3	US-09-949-016-16514	Sequence 16514, A
C 44	13.4	89.3	784019	3	US-09-949-016-14033	Sequence 14033, A
C 45	13.4	89.3	828152	3	US-09-949-016-12777	Sequence 12777, A

ALIGNMENTS

RESULT 1

US-09-248-796A-3043

; Sequence 3043, Application US/09248796A

; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 3043

; LENGTH: 1323

; TYPE: DNA

; ORGANISM: Candida albicans

; US-09-248-796A-3043

Query Match 93.3%; Score 14; DB 3; Length 1323;

Best Local Similarity 100.0%; Pred. No. 2.8e+02;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCCCGGTACC 14

Db 1025 CTGCCCGGTACC 1038

RESULT 2

US-09-221-017B-510

; Sequence 510, Application US/09221017B

; Patent No. 6444799

; GENERAL INFORMATION:

; APPLICANT: Ross, Bruce C.

; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF

; NUMBER OF SEQUENCES: 1120

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORRISON & FOERSTER

; STREET: 755 PAGE MILL ROAD

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304-1018

; COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSQL for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P11182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P11546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P22911
FILING DATE: 09-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Monroy, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 510:
SEQUENCE CHARACTERISTICS:
LENGTH: 7319 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...7319
US-09-221-017B-510

Query Match 93.3%; Score 14; DB 3; Length 7319;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCCCCACGGTACC 14
|||
Db 3741 CTGCCCCACGGTACC 3754

RESULT 3
US-08-707-237A-5
Sequence 5, Application US/08707237A
Patent No. 5830713
GENERAL INFORMATION:
APPLICANT: Ferrari, Franco A.
APPLICANT: Capello, Joseph
APPLICANT: Crissman, John W.
APPLICANT: Dorman, Mary A.
TITLE OF INVENTION: METHODS FOR PREPARING SYNTHETIC
TITLE OF INVENTION: REPETITIVE DNA
NUMBER OF SEQUENCES: 108
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/707,237A
FILING DATE: 03-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/175,155
FILING DATE: 29-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/053,049
FILING DATE: 22-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/609,716
FILING DATE: 06-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/269,429
FILING DATE: 09-NOV-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/114,618
FILING DATE: 29-OCT-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/927,258
FILING DATE: 04-NOV-1986
ATTORNEY/AGENT INFORMATION:
NAME: Trecartin, Richard P.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-55186-10/WHO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 63 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-707-237A-5

Query Match 89.3%; Score 13.4; DB 2; Length 63;
Best Local Similarity 93.3%; Pred. No. 5.8e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTGCCCCACGGTACCT 15
|||
Db 26 CTACCCACGGTACCT 40

RESULT 4
US-08-882-164D-33/c
Sequence 33, Application US/08882164D
Patent No. 6306624
GENERAL INFORMATION:
APPLICANT: Petkovich, P. Martin, White, Jay A.,
APPLICANT: Beckett, Barbara R., Jones, Glenville
TITLE OF INVENTION: Retinoid Metabolizing Protein
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Blake, Cassels & Graydon
STREET: Box 25, Commerce Court West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5L 1A9
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage
COMPUTER: COMPAQ, IBM PC compatible
OPERATING SYSTEM: MS-DOS 5.1
SOFTWARE: WORD PERFECT

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 15:49:05 ; Search time 785.792 Seconds
(without alignments)
1157.425 Million cell updates/sec

Title: US-10-754-446-7

Perfect score: 16

Sequence: 1 agaccaggccacat 16

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sts.*
11: gb_ay.*
12: gb_un.*
13: gb_vi.*
14: gb_hlg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	16	100.0	536	10 BV401493	BV401493 S229P6332
C 2	16	100.0	579	8 AF305572S5	AF305576 Homo sapi
C 3	16	100.0	659	10 BV230130	BV230130 S233P6354
C 4	16	100.0	661	10 G93242	G93242 S208P6433RE
C 5	16	100.0	728	10 BV524331	BV524331 G591P6065
C 6	16	100.0	925	10 BV526258	BV526258 G591P6047
C 7	16	100.0	1740	6 AX280021	AX280021 Sequence
C 8	16	100.0	1829	6 AQ719994	AQ719994 Sequence
C 9	16	100.0	2004	8 AF249319	AF249319 Homo sapi
C 10	16	100.0	2037	8 HSA293970	AJ293970 Homo sapi
C 11	16	100.0	2049	8 AK026102	AK026102 Homo sapi
C 12	16	100.0	2051	8 AF287269	AF287269 Homo sapi
C 13	16	100.0	2052	8 AX083508	AX083508 Sequence
C 14	16	100.0	2063	8 AK222673	AK222673 Homo sapi
C 15	16	100.0	2065	8 AB125179	AB125179 Macaca fa
C 16	16	100.0	2078	8 BC005149	BC005149 Homo sapi
C 17	16	100.0	2094	6 BD233734	BD233734 31 human
C 18	16	100.0	2095	6 AX280019	AX280019 Sequence

19	16	100.0	2272	8 HSA293659	AJ293659 Homo sapi
C 20	16	100.0	12300	1 AE005887	AE005887 Caulobact
21	16	100.0	13270	8 AF287270	AF287270 Homo sapi
22	16	100.0	47196	14 BX546462_4	Continuation (5 of
C 23	16	100.0	49264	14 AC164672	AC164672 Bos tauru
C 24	16	100.0	128501	8 AC012614	AC012614 Homo sapi
C 25	16	100.0	150909	9 AC116811	AC116811 Mus muscu
C 26	16	100.0	151130	9 AC127432	AC127432 Mus muscu
C 27	16	100.0	155645	14 AC021153	AC021153 Homo sapi
C 28	16	100.0	157952	8 AF003691	AF003691 Homo sapi
C 29	16	100.0	158063	8 AP001046	AP001046 Homo sapi
C 30	16	100.0	165994	9 AC129207	AC129207 Mus muscu
31	16	100.0	173081	9 AC159338	AC159338 Mus muscu
32	16	100.0	173126	8 AC008878	AC008878 Homo sapi
C 33	16	100.0	175177	9 AC124383	AC124383 Mus muscu
34	16	100.0	179556	14 AC018734	AC018734 Homo sapi
35	16	100.0	180770	14 AC153861	AC153861 Mus muscu
36	16	100.0	188593	8 AC016877	AC016877 Homo sapi
37	16	100.0	194042	9 AC153850	AC153850 Mus muscu
C 38	16	100.0	194364	14 AC164574	AC164574 Mus muscu
C 39	16	100.0	208898	14 AC164628	AC164628 Mus muscu
C 40	16	100.0	210408	8 AC010608	AC010608 Homo sapi
41	16	100.0	213826	9 AL845298	AL845298 Mouse DNA
42	16	100.0	222368	14 AC095138	AC095138 Rattus no
C 43	16	100.0	225535	14 AC111381	AC111381 Rattus no
C 44	16	100.0	232747	14 AC116259	AC116259 Rattus no
45	16	100.0	253693	14 AC131839	AC131839 Rattus no

ALIGNMENTS

RESULT 1
BV401493/c

LOCUS

DEFINITION BV229P6332R8.T0 GermanShepherd Canis familiaris STS genomic,
sequence tagged site.

ACCESSION BV401493

VERSION BV401493.1

KEYWORDS GI:57775274

SOURCE STS.

ORGANISM Canis familiaris (dog)

Canis familiaris

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;

Canis.

REFERENCE 1 (bases 1 to 536)

AUTHORS Lindblad-Toh, K.

TITLE The genome sequence of Canis familiaris

JOURNAL Unpublished (2004)

COMMENT

Contact: Kerstin Lindblad-Toh
Whitehead Institute for Biomedical Research, Center for Genome
Research
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172521477
Fax: 6172580903
Email: kersli@genome.wi.mit.edu
Primer A: No sequence submitted
Primer B: No sequence submitted
STS size: 536
Protocol:

WGS-discovery (WGS):
Paired-end low-coverage whole genome shotgun reads were generated
from 9 breeds
(German Shepherd, Rottweiler, Bedlington Terrier, Beagle, Labrador
Retriever, English
Shepherd, Italian Greyhound, Alaskan Malamute and the Portuguese
Water Dog -100,000 each)
and five other canids (Chinese, Alaskan, Indian and Spanish Gray
Wolf as well as the
Californian Coyote).
The WGS reads were placed uniquely on the CanFam1.0 boxer assembly
and SNP detection was

carried out by SSAHA-SNP. 863872 reads were annotated as STSs and 485941 SNPs were annotated with alleles from the boxer and the breed or canid from which the particular read came. The validation rate for these SNPs was estimated at approximately 98%.

WGA-discovery (WGA) of Boxer/Poodle SNPs: A second set of SNPs was generated using a similar methodology except that the contigs from the 1.5x poodle assembly (Kirkness 2003) were used instead of WGS reads. Since this sequence lacked base quality scores, arbitrary quality scores of phred 40 were assigned before the poodle sequence was placed uniquely on the CanFam1.0 boxer assembly and SNP detection was carried out by SSAHA-SNP. 1637780 SNPs were annotated with alleles from the boxer and the breed or canid from which the particular read came. The validation rate for these SNPs was estimated at approximately 98%.

Internal-WGA-discovery (I-WGA): A third set of SNPs were discovered by comparing reads in the WGA assembly. SNPs were defined as mismatch positions that had a base quality of ≥ 30 on both reads in a region that aligned without gaps, and with at most one additional mismatch in the ten flanking bases. For each allele, at least one additional read had to confirm it. 731476 SNPs were annotated with alleles between the two boxer alleles. The validation rate for these SNPs was estimated at approximately 98%.

FEATURES
source
1. .536
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="GermanShepherd"
/db_xref="taxon:9615"
/map="* 8 22-462 74954487-74954047"
/clone_lib="GermanShepherd"
<1. .536

STS
ORIGIN

Query Match 100.0%; Score 16; DB 10; Length 536;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGACCCAGGCCACAT 16
Db 468 AGACCCAGGCCACAT 453

RESULT 2
AF305572S5
LOCUS AF305572S5 579 bp DNA linear PRI 26-DEC-2000
DEFINITION Homo sapiens mucopolipin 1 (MCOLN1) gene, exons 6 and 7.
ACCESSION AF305576
VERSION AF305576.1 GI:11991202
KEYWORDS
SEGMENT 5 of 8
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 579)
Bargal, R., Avidan, N., Ben-Asher, E., Olender, Z., Zeigler, M., Frumkin, A., Raas-Rothschild, A., Glusman, G., Lancet, D. and Bach, G.
TITLE Identification of the gene causing mucopolipidosis type IV
JOURNAL Nat. Genet. 26 (1), 118-123 (2000)
PUBMED 10973263

REFERENCE
AUTHORS 2 (bases 1 to 579)
Bargal, R., Avidan, N., Ben-Asher, E., Olender, A., Zeigler, M., Frumkin, A., Raas-Rothschild, A., Glusman, G., Lancet, D. and Bach, G.

TITLE Direct Submission
JOURNAL Submitted (13-SEP-2000) Molecular Genetics, The Weizmann Institute of Science, P. O. Box 26, Rehovot 76100, Israel

FEATURES
source
1. .579
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
129_225
/gene="MCOLN1"
/number=6
423. .522
/gene="MCOLN1"
/number=7

exon
exon

ORIGIN

Query Match 100.0%; Score 16; DB 8; Length 579;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGACCCAGGCCACAT 16
Db 472 AGACCCAGGCCACAT 487

RESULT 3
BV230130/c
LOCUS BV230130 659 bp DNA linear STS 19-JAN-2005
DEFINITION S233P6354FD9.T0 LabradorRetriever Canis familiaris STS genomic, sequence tagged site.
ACCESSION BV230130
VERSION BV230130.1 GI:57292672
KEYWORDS STS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae; Canis.

REFERENCE
AUTHORS 1 (bases 1 to 659)
Lindblad-Toh, K.
TITLE The genome sequence of Canis familiaris
JOURNAL Unpublished (2004)
COMMENT Contact: Kerstin Lindblad-Toh
Whitehead Institute for Biomedical Research, Center for Genome Research
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172521477
Fax: 6172580903
Email: Kersti@genome.wi.mit.edu
Primer A: No sequence submitted
Primer B: No sequence submitted
STS size: 659
Protocol: WGS-discovery (WGS):
Paired-end low-coverage whole genome shotgun reads were generated from 9 breeds (German Shepherd, Rottweiler, Bedlington Terrier, Beagle, Labrador Retriever, English Shepherd, Italian Greyhound, Alaskan Malamute and the Portuguese Water Dog -100,000 each) and five other canids (Chinese, Alaskan, Indian and Spanish Gray Wolf as well as the Californian Coyote).
The WGS reads were placed uniquely on the CanFam1.0 boxer assembly and SNP detection was carried out by SSAHA-SNP. 863872 reads were annotated as STSs and 485941 SNPs were annotated with alleles from the boxer and the breed or canid from which the particular read came. The validation rate for these SNPs was estimated at approximately 98%.

WGA-discovery (WGA) of Boxer/Poodle SNPs:

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 15:55:50 ; Search time 1789.31 Seconds
(without alignments)
418.369 Million cell updates/sec

Title: US-10-754-446-7
Perfect score: 16
Sequence: 1 agaccagggccacat 16

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_hic.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_est7.*
9: gb_gss1.*
10: gb_gss2.*
11: gb_gss3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	100.0	278	9	AZ491094
2	16	100.0	291	8	T29948
3	16	100.0	316	8	F07345
4	16	100.0	324	1	AW486159
5	16	100.0	331	8	Z43611
6	16	100.0	333	8	T35364
7	16	100.0	350	8	Z43604
8	16	100.0	371	1	AW356454
9	16	100.0	387	1	AW425724
10	16	100.0	387	2	BE478501
11	16	100.0	404	9	AZ847535
12	16	100.0	442	1	AW785343
13	16	100.0	466	9	AZ073212
14	16	100.0	478	2	BG383183
15	16	100.0	500	2	BF443224
16	16	100.0	510	6	CA889269
17	16	100.0	540	7	CJ017244
18	16	100.0	578	9	AZ286475
19	16	100.0	593	7	CN788730
20	16	100.0	740	6	CA489568
21	16	100.0	772	7	CK465307
22	16	100.0	775	10	CG740895

c	23	16	100.0	777	7	CK467406
c	24	16	100.0	779	7	CK468501
c	25	16	100.0	817	2	B1252802
c	26	16	100.0	826	2	BG913075
c	27	16	100.0	845	2	BG176074
c	28	16	100.0	862	7	CK469557
c	29	16	100.0	868	3	B1909885
c	30	16	100.0	872	10	CZ363216
c	31	16	100.0	873	6	CD558014
c	32	16	100.0	878	5	BQ956452
c	33	16	100.0	901	5	BX433310
c	34	16	100.0	961	10	CZ363668
c	35	16	100.0	979	5	BQ918296
c	36	16	100.0	993	2	BG739714
c	37	16	100.0	1056	11	CNS037M4
c	38	16	100.0	1078	5	BX362786
c	39	16	100.0	1273	4	CR616477
c	40	16	100.0	1712	10	AY410284
c	41	16	100.0	1743	10	AY410283
c	42	16	100.0	1996	4	CR622331
c	43	16	100.0	2108	4	AK052519
c	44	15.6	97.5	961	5	BX397924
c	45	15.2	95.0	1026	5	BX364018

ALIGNMENTS

RESULT 1
LOCUS AZ491094 278 bp DNA linear GSS 05-OCT-2000
DEFINITION IM0324P22F Mouse 10kb plasmid UUGCLM library Mus musculus genomic clone UUGCLM0324P22 F, genomic survey sequence.
ACCESSION AZ491094
VERSION AZ491094.1 GI:10662462
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 278)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0324 row: P column: 22
Seq primer: GGTGTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 278.
Location/Qualifiers
1. 278
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCLM0324P22"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, P-"
/clone_lib="Mouse 10kb plasmid library"
/note="Vector: PWD42hv; Purified genomic DNA from M.

FEATURES
source

musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gii4732114[gblAF129072.1]), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 100.0%; Score 16; DB 9; Length 278;
Best Local Similarity 100.0%; Pred. No. 4e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGACCCAGGCCACAT 16
|||||
DB 129 AGACCCAGGCCACAT 144

RESULT 2

T29948
LOCUS EST100452 Human Pancreas Homo sapiens cDNA 5' end similar to None,
DEFINITION mRNA sequence.

ACCESSION T29948.1 GI:612046
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE

1 (bases 1 to 291)
Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult, C.J., Lee, N., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Chiu, M.-W., Clayton, R.A., Cline, R.T., Cotton, M.D., Earle-Hughes, J., Fine, L.D., FitzGerald, L.M., FitzHugh, W.M., Fritchman, J.L., Geoghagen, N.S.M., Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle Jr, P.S., Kelley, J.M., Kline, K.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Balances, R.F., McDonald, L.A., Nguyen, D.T., Pellegrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudak, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarek, D.P., Cao, L., Cepeda, N.A., Coleman, T.A., Collins, E.-J., Dimke, D., Feng, P., Ferrie, A., Fischer, C., Hastings, G.A., He, W.-W., Hu, J.-S., Greene, J.M., Gruber, J., Hudson, P., Kim, A., Kozak, D.L., Kunsch, C., Ji, H., Li, H., Meisner, P.S., Olsen, H., Raymond, L., Wei, Y.-F., Wing, J., Xu, C., Yu, G.-L., Ruben, S.M., Dillon, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.
Initial Assessment of Human Gene Diversity and Expression Patterns
Based Upon 83 Million Basepairs of cDNA Sequence
Nature 377, 3-174 (1995)

TITLE
JOURNAL
PUBMED
COMMENT

Other_ESTs: THC11326
Contact: Venter, JC
The Institute for Genomic Research
932 Clopper Rd, Gaithersburg, MD 20878
Tel: 3018695056
Fax: 3018693423

Email: tdbinfo@db.tigr.org
For clone availability, additional sequence and expression information related to this EST, please contact the TIGR Database

(tdbinfo@db.tigr.org)
Seq primer: M13 Reverse.
Location/Qualifiers
1. .291
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):107961"
/db_xref="taxon:9606"
/clone_lib="Human Pancreas"
/note="Organ: pancreas"

FEATURES
source

ORIGIN

Query Match 100.0%; Score 16; DB 8; Length 291;
Best Local Similarity 100.0%; Pred. No. 4e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGACCCAGGCCACAT 16
|||||
DB 186 AGACCCAGGCCACAT 201

RESULT 3

F07345
LOCUS HSC22G041 normalized infant brain cDNA Homo sapiens CDNA clone
DEFINITION c-22g04, mRNA sequence.

ACCESSION F07345.1 GI:672999
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE

1 (bases 1 to 316)
Auffray, C., Behar, G., Bois, F., Bouchier, C., da Silva, C., Devignes, M.D., Duprat, S., Houlgatte, R., Jumeau, M.N., Lamy, B., Lorenzo, F., Mitchell, H., Mariage-Samson, R., Pietu, G., Pouliot, Y., Sebastiani-Kabaktchis, C., and Tessier, A.
IMAGE: molecular integration of the analysis of the human genome and its expression
C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
7757816

COMMENT

Contact: Genethon
Genexpress-Genethon
Genethon Centre de recherche sur le Genome Humain
1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
Tel: 33169472800
Fax: 33160778698
Email: genexpress@genethon.fr
Single read
Genexpress_library_idt: C; Genexpress_sequence_idt: ylc-22g04
Seq primer: (-21)M13 universal.
Location/Qualifiers

FEATURES
source

1. .316
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="c-22g04"
/sex="Female"
/tissue_type="total brain"
/dev_stage="3 months old"
/clone_lib="normalized infant brain cDNA"
/notes="Organ: brain; Vector: lafmid BA; Site_1: HindIII; Site_2: NotI; sex:Female; dev_stage=3 months old; isolate=muscular atrophy patient; tissue_type=total brain; total mRNA was oligo-(dT) primed and directionally cloned 5' -> 3' into the HindIII -> NotI sites of the lafmid BA vector. Clone library from B. Soares, Psychiatry Dept. Columbia University, USA. Normalization_method: Bento Soares, P.N.A.S in press"

ORIGIN

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 14:15:14 ; Search time 217.088 Seconds
(without alignments)
491.207 Million cell updates/sec

Title: US-10-754-446-7

Perfect score: 16

Sequence: 1 agaccaggccacat 16

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16	100.0	16	14	AEB28963
2	16	100.0	1619	4	AAL1022 Human pol
3	16	100.0	1740	6	AAL1700 Human TRP
4	16	100.0	1740	6	ABL40755 Human TLC
5	16	100.0	1741	4	AAL59236 Human pol
6	16	100.0	1743	9	ADB84284 Human muc
7	16	100.0	1743	14	ADV66231 TRP-like
8	16	100.0	2051	13	ADP25150 PRO polyp
9	16	100.0	2052	4	AAF81753 Human mem
10	16	100.0	2092	5	AAL59236 Human pol
11	16	100.0	2092	6	ABL90358 Human pol
12	16	100.0	2094	3	AA339067 Human sec
13	16	100.0	2095	6	AA171699 Human TRP
14	16	100.0	2095	6	ABL40754 Human TLC
15	16	100.0	2095	14	ADV66229 TRP-like
16	16	100.0	2140	12	ADQ24925 Human sof
17	16	100.0	11365	4	AAK73827 Human imm
18	16	100.0	13270	9	ADB84283 Human muc
19	16	100.0	20046	4	AAK73826 Human imm

20	16	100.0	340449	8	AAL52198	AAL52198 Human sec
c 21	15	93.8	384	6	ABN15902	ABN15902 Human ORF
c 22	15	93.8	449	11	ADT97636	ADT97636 Colon can
c 23	15	93.8	449	11	ADX44118	ADX44118 Human CDN
c 24	15	93.8	471	13	ADU10604	ADU10604 Solid tum
c 25	15	93.8	502	4	AAK61151	AAK61151 Human imm
c 26	15	93.8	514	13	ADQ49858	ADQ49858 Novel can
c 27	15	93.8	625	4	AAH05091	AAH05091 Human CDN
c 28	15	93.8	819	3	AACT9832	AACT9832 Human sec
c 29	15	93.8	882	4	AAH34308	AAH34308 Human col
c 30	15	93.8	882	6	ABL90324	ABL90324 Human pol
c 31	15	93.8	943	4	AAK79476	AAK79476 Human imm
c 32	15	93.8	943	4	AAK79477	AAK79477 Human imm
c 33	15	93.8	1422	5	AAH83077	AAH83077 DNA encod
c 34	15	93.8	1468	4	AAH15702	AAH15702 Human CDN
c 35	15	93.8	1957	10	ADES9228	ADES9228 Human gen
c 36	15	93.8	1957	10	ADES9225	ADES9225 Human gen
c 37	15	93.8	2000	11	ACL37128	ACL37128 Rice stre
c 38	15	93.8	2004	10	ADB63068	ADB63068 Human CDN
c 39	15	93.8	3111	12	ADN38521	ADN38521 Novel hum
c 40	15	93.8	3444	11	ADL22622	ADL22622 Human dis
c 41	15	93.8	3600	8	ADA83759	ADA83759 Human KIA
c 42	15	93.8	3600	12	ADI24494	ADI24494 Human mod
c 43	15	93.8	3600	14	ADX83158	ADX83158 Human TEG
c 44	15	93.8	3600	14	ADX44432	ADX44432 Human col
c 45	15	93.8	3600	14	ADY39343	ADY39343 Human col

ALIGNMENTS

RESULT 1

AEB28963
ID AEB28963 standard; DNA; 16 BP.
XX
AC AEB28963;
XX
DT 22-SEP-2005 (first entry)
XX
DE Human MCOLN1 gene probe DEL SEQ ID NO:7.
XX
KW mucopolipidosis IV; high throughput screening; mucolipin-1; MCOLN1; probe;
KW ss.
XX
OS Homo sapiens.
XX
PN US2005153300-A1.
XX
PD 14-JUL-2005.
XX
PF 09-JAN-2004; 2004US-00754446.
XX
PR 09-JAN-2004; 2004US-00754446.
XX
(QUES-) QUEST DIAGNOSTICS INC.
XX
Sun W, Hantash F;
WPI; 2005-521160/53.

Diagnosing mucopolipidosis (ML) IV mutant sequence in nucleic acids, by amplifying the nucleic acid, detecting amplified product with labeled oligonucleotide probes via a change in fluorescence which indicates the presence of an ML IV mutant.

Claim 3; SEQ ID NO 7; 15pp; English.

The invention relates to a method (M1) for detecting the presence of mucopolipidosis IV (ML IV) mutant sequences in a nucleic acid. (M1) involves contacting the nucleic acid with oligonucleotide primers and probes, conducting amplification by temperature cycling and monitoring the accumulation of amplified nucleic acid by detecting an increase in donor fluorophore fluorescence or decrease in acceptor fluorophore

CC fluorescence, which indicates the presence of ML IV mutant sequence. Also
 CC described: (1) an oligonucleotide (ON1) comprising a sequence
 CC complementary to the coding or non-coding strand of the mucolin-1
 CC (MCOLN1) gene; and (2) a kit (KI) for amplifying sequences of ML IV
 CC sequences, comprising ON1. (M1) is useful for determining the presence of
 CC ML IV mutant sequence in a nucleic acid, and/or wild-type sequences. (M1)
 CC is useful for genetic counseling of individuals at risk for ML IV or at
 CC risk to having offspring afflicted with ML IV. The MCOLN1 gene is located
 CC on chromosome 19p13. The present sequence represents a probe for the
 CC MCOLN1 gene, which is used in the exemplification of the present
 CC invention.

XX
 SQ Sequence 16 BP; 5 A; 7 C; 3 G; 1 T; 0 U; 0 Other;
 Query Match 100.0%; Score 16; DB 14; Length 16;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02; Mismatches 0; Indels 0; Gaps 0;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGACCCAGGCCACAT 16
 |||||
 Db 1 AGACCCAGGCCACAT 16
 |||||

RESULT 2
 AAI61022/c
 ID AAI61022 standard; cDNA; 1619 BP.
 XX
 AC AAI61022;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 5011.
 XX
 KW Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US034263.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang J, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI; 2001-442253/47.
 DR P-PSDB; AAM41866.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX
 PS Claim 1; SEQ ID NO 5011; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the

CC encoded polypeptides (AAM38642-AAM42213) with neurotropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Actin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemia and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification

XX
 SQ Sequence 1619 BP; 350 A; 450 C; 518 G; 301 T; 0 U; 0 Other;
 Query Match 100.0%; Score 16; DB 4; Length 1619;
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGACCCAGGCCACAT 16
 |||||
 Db 1122 AGACCCAGGCCACAT 1107
 |||||

RESULT 3
 AAI71700
 ID AAI71700 standard; cDNA; 1740 BP.
 XX
 AC AAI71700;
 XX
 DT 29-JAN-2002 (first entry)
 XX
 DE Human TRP-like calcium channel TLCC-2 coding sequence #2.
 XX
 KW Human; TLCC-2; TRP-like calcium channel; membrane excitability;
 KW nociception; neurotropic; neuroprotective; antiparkinsonian; cytostatic;
 KW hypotensive; antidepressant; analgesic; anticonvulsant; tranquiliser;
 KW Parkinson's disease; Huntington's disease; multiple sclerosis;
 KW Gilles de la Tourette's syndrome; autonomic function disorder; cancer;
 KW neuroleptic; gene therapy; Alzheimer's disease; CNS disorder; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1740
 FT /*tag= a
 FT /product= "TLCC-2"
 FT /partial
 XX
 WO200177331-A1.
 XX
 PD 18-OCT-2001.
 XX
 PF 06-APR-2001; 2001WO-US011442.
 XX
 PR 07-APR-2000; 2000US-00544797.
 XX
 PA (MILL-) MILLENIUM PHARM INC.
 XX
 PI Curtis RAJ, Silos-Santiago I;
 XX
 DR WPI; 2002-010913/01.
 DR P-PSDB; AAM51858.
 XX
 PT Novel isolated human transient receptor potential-like calcium channel
 PT protein-2 useful for treating Alzheimer's disease, depression, amnesia,
 PT pain disorder, and cancer.
 XX
 PS Claim 1; Fig 1; 148pp; English.
 XX
 CC The present invention relates to the protein and coding sequences of

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 16:41:33 ; Search time 386.432 Seconds
(without alignments)
342.389 Million cell updates/sec

Title: US-10-754-446-7

Perfect score: 16

Sequence: 1 agaccaggccacat 16

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	100.0	16	9 US-10-754-446-7	Sequence 7, Appli
2	16	100.0	1400	10 US-11-060-756-2850	Sequence 2850, Ap
3	16	100.0	1400	10 US-11-060-756-7122	Sequence 7122, Ap
4	16	100.0	1740	3 US-09-828-466-3	Sequence 3, Appli
5	16	100.0	1740	5 US-10-103-458-3	Sequence 3, Appli
6	16	100.0	1743	8 US-10-782-695-6	Sequence 6, Appli
7	16	100.0	2051	3 US-09-851-494B-2	Sequence 2, Appli
8	16	100.0	2052	3 US-09-965-529-50	Sequence 50, Appli
9	16	100.0	2052	3 US-09-969-680A-50	Sequence 50, Appli
10	16	100.0	2052	10 US-11-048-692-50	Sequence 50, Appli
11	16	100.0	2092	6 US-10-284-237-920	Sequence 920, App
12	16	100.0	2092	9 US-10-450-763-8078	Sequence 8078, Ap
13	16	100.0	2094	3 US-09-820-893-26	Sequence 26, Appli
14	16	100.0	2094	7 US-10-607-565-26	Sequence 1, Appli
15	16	100.0	2095	3 US-09-828-466-1	Sequence 1, Appli
16	16	100.0	2095	5 US-10-103-458-1	Sequence 4, Appli
17	16	100.0	2095	8 US-10-782-695-4	Sequence 7445, Ap
18	16	100.0	2140	8 US-10-723-860-7745	Sequence 1, Appli
19	16	100.0	13270	3 US-09-851-494B-1	Sequence 3, Appli
20	16	100.0	34049	3 US-09-903-582-3	Sequence 3, Appli
21	15	93.8	342	7 US-10-437-963-97094	Sequence 97094, A
22	15	93.8	445	5 US-10-066-543-3155	Sequence 3155, Ap
23	15	93.8	552	7 US-10-767-701-22677	Sequence 22677, A

24	15	93.8	558	4	US-09-925-065A-531750	Sequence 531750,
c 25	15	93.8	596	5	US-10-027-632-220987	Sequence 220987,
c 26	15	93.8	609	4	US-10-027-632-220987	Sequence 220987,
c 27	15	93.8	614	4	US-09-925-065A-502281	Sequence 502281,
c 28	15	93.8	614	4	US-09-925-065A-73411	Sequence 73411, A
c 29	15	93.8	614	4	US-09-925-065A-73412	Sequence 73412, A
c 30	15	93.8	614	4	US-09-925-065A-73413	Sequence 73413, A
c 31	15	93.8	627	4	US-09-925-065A-647660	Sequence 647660,
c 32	15	93.8	675	4	US-09-925-065A-703658	Sequence 703658,
c 33	15	93.8	675	4	US-09-925-065A-703659	Sequence 703659,
c 34	15	93.8	675	4	US-09-925-065A-703660	Sequence 703660,
c 35	15	93.8	678	8	US-10-425-115-77806	Sequence 77806, A
c 36	15	93.8	882	5	US-10-106-698-1400	Sequence 1400, Ap
c 37	15	93.8	882	6	US-10-284-237-886	Sequence 886, App
c 38	15	93.8	1422	9	US-10-450-763-18881	Sequence 18881, A
c 39	15	93.8	2004	6	US-10-104-047-1222	Sequence 1222, Ap
c 40	15	93.8	3600	6	US-10-157-031-41	Sequence 41, Appl
c 41	15	93.8	3600	8	US-10-651-237-52	Sequence 52, Appl
c 42	15	93.8	3600	8	US-10-782-413-52	Sequence 52, Appl
c 43	15	93.8	52510	5	US-10-087-192-1006	Sequence 1006, Ap
c 44	15	93.8	59747	7	US-10-450-826-116	Sequence 116, App
c 45	15	93.8	197775	5	US-10-087-192-853	Sequence 853, App

ALIGNMENTS

RESULT 1
US-10-754-446-7
; Sequence 7, Application US/10754446
; Publication No. US20050153300A1
; GENERAL INFORMATION:
; APPLICANT: HANTASH, FERAS
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DETECTION OF
; FILE REFERENCE: 034827-2301
; CURRENT APPLICATION NUMBER: US/10754.446
; CURRENT FILING DATE: 2004-01-09
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 7
; LENGTH: 16
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Probe
US-10-754-446-7

Query Match 100.0%; Score 16; DB 9; Length 16;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGACCCAGGCCACAT 16
| | | | | | | | | | | | | | | |
Db 1 AGACCCAGGCCACAT 16

RESULT 2
US-11-060-756-2850
; Sequence 2850, Application US/11060756
; Publication No. US20050221354A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060.756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2850

```
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-060-756-2850

Query Match      100.0%; Score 16; DB 10; Length 1400;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGACCCAGGCCACAT 16
    |||||
Db 301 AGACCCAGGCCACAT 316

RESULT 3
US-11-060-756-7122
; Sequence 7122, Application US/11060756
; Publication No. US20050221354A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; TITLE OF INVENTION: Target Genes
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060,756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7122
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-060-756-7122

Query Match      100.0%; Score 16; DB 10; Length 1400;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGACCCAGGCCACAT 16
    |||||
Db 301 AGACCCAGGCCACAT 316

RESULT 4
US-09-828-466-3
; Sequence 3, Application US/09828466
; Patent No. US20020035056A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Silos-Santiago, Immaculada
; TITLE OF INVENTION: 54420, A NOVEL HUMAN CALCIUM CHANNEL
; FILE REFERENCE: MNI-125CP
; CURRENT APPLICATION NUMBER: US/09/828,466
; CURRENT FILING DATE: 2001-04-06
; PRIOR FILING DATE: 2000-04-07
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1740)
US-09-828-466-3

Query Match      100.0%; Score 16; DB 3; Length 1740;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 16; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGACCCAGGCCACAT 16
    |||||
; Remaining Prior Application data removed - See File Wrapper or PALM.
```

```
Db 827 AGACCCAGGCCACAT 842

RESULT 5
US-10-103-458-3
; Sequence 3, Application US/10103458
; Publication No. US20020197680A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; TITLE OF INVENTION: 54420, A NOVEL HUMAN CALCIUM CHANNEL
; FILE REFERENCE: MNI-125
; CURRENT APPLICATION NUMBER: US/10/103,458
; CURRENT FILING DATE: 2002-03-22
; PRIOR FILING DATE: US/09/544,797
; PRIOR FILING DATE: PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1740)
US-10-103-458-3

Query Match      100.0%; Score 16; DB 5; Length 1740;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGACCCAGGCCACAT 16
    |||||
Db 827 AGACCCAGGCCACAT 842

RESULT 6
US-10-782-695-6
; Sequence 6, Application US/10782695
; Publication No. US20040248160A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Tsai, Fong-Ying
; APPLICANT: Hodge, Martin R.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: NOVEL 14275, 54420, 8797, 27439, 68730,
; FILE REFERENCE: MPI04-0020NNIM
; CURRENT APPLICATION NUMBER: US/10/782,695
; CURRENT FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: US 09/945,254
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,829
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 09/945,301
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,301
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 10/007,399
; PRIOR FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: US 09/390,039
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: US 09/146,416
; PRIOR FILING DATE: 1998-09-03
; PRIOR APPLICATION NUMBER: US 10/024,036
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/258,222
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 10/103,458
; PRIOR FILING DATE: 2002-03-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 16:59:57 ; Search time 196.864 Seconds
(without alignments)
65.779 Million cell updates/sec

Title: US-10-754-446-7

Perfect score: 16

Sequence: 1 agaccaggccacat 16

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 6038814 seqs, 404674181 residues

Total number of hits satisfying chosen parameters: 12077628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA New:*

- 1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
- 9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
- 10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	100.0	172147	7 US-11-112-908-22	Sequence 22, Appl
2	16	100.0	188682	7 US-11-112-908-23	Sequence 23, Appl
3	15	93.8	22	6 US-10-310-914A-1360219	Sequence 1360219,
4	15	93.8	25	6 US-10-310-914A-1360184	Sequence 1360184,
5	15	93.8	403278	6 US-10-995-561-13421	Sequence 13421, A
6	14.4	90.0	19	8 US-11-101-244-97619	Sequence 97619, A
7	14.4	90.0	19	8 US-11-101-244-97633	Sequence 97633, A
8	14.4	90.0	19	9 US-11-083-784-97619	Sequence 97619, A
9	14.4	90.0	19	9 US-11-083-784-97633	Sequence 97633, A
10	14.4	90.0	791	6 US-10-750-185-62019	Sequence 62019, A
11	14.4	90.0	791	6 US-10-750-623-62019	Sequence 62019, A
12	14.4	90.0	1214	7 US-11-112-908-134	Sequence 134, App
13	14.4	90.0	1889	6 US-10-821-234-37	Sequence 37, Appl
14	14.4	90.0	2215	7 US-11-112-908-144	Sequence 144, App
15	14.4	90.0	2453	6 US-10-750-185-36539	Sequence 36539, A
16	14.4	90.0	2453	6 US-10-750-623-36539	Sequence 36539, A
17	14.4	90.0	44848	7 US-11-106-672A-42	Sequence 42, Appl
18	14.4	90.0	54767	6 US-10-995-561-13357	Sequence 13357, A
19	14.4	90.0	158410	7 US-11-121-086-46	Sequence 46, Appl
20	14.4	90.0	164810	7 US-11-121-086-46	Sequence 4, Appl
21	14.4	90.0	180862	7 US-11-112-908-40	Sequence 40, Appl
22	14.4	90.0	189539	7 US-11-121-086-16	Sequence 16, Appl
23	14	87.5	201	7 US-11-124-368A-4143	Sequence 4143, Ap

24	14	87.5	496	6 US-10-750-185-37291	Sequence 37291, A
25	14	87.5	496	6 US-10-750-623-37291	Sequence 37291, A
26	14	87.5	600	7 US-11-136-527-6328	Sequence 6328, Ap
c 27	14	87.5	1038	6 US-10-512-109-16	Sequence 16, Appl
c 28	14	87.5	1223	6 US-10-512-109-3	Sequence 3, Appl
c 29	14	87.5	1223	6 US-10-512-109-14	Sequence 14, Appl
c 30	14	87.5	1816	6 US-10-750-185-29826	Sequence 29826, A
c 31	14	87.5	1816	6 US-10-750-623-29826	Sequence 29826, A
c 32	14	87.5	2378	6 US-10-821-234-117	Sequence 117, App
c 33	14	87.5	3884	6 US-10-131-826A-145	Sequence 145, App
c 34	14	87.5	10705	7 US-11-136-527-2232	Sequence 2232, Ap
c 35	14	87.5	20945	6 US-10-995-561-13463	Sequence 13463, A
c 36	14	87.5	23082	6 US-10-995-561-13457	Sequence 13457, A
c 37	14	87.5	27615	7 US-11-136-527-514	Sequence 514, App
c 38	14	87.5	94510	6 US-10-995-561-13332	Sequence 13332, A
c 39	14	87.5	103660	6 US-10-995-561-13253	Sequence 13253, A
c 40	14	87.5	197096	7 US-11-121-086-107	Sequence 1165996,
c 41	13.4	83.8	19	6 US-10-310-914A-1165996	Sequence 107, App
c 42	13.4	83.8	19	8 US-11-101-244-798463	Sequence 798463,
c 43	13.4	83.8	19	9 US-11-083-784-798463	Sequence 798463,
c 44	13.4	83.8	21	6 US-10-310-914A-1130386	Sequence 1130386,
c 45	13.4	83.8	23	6 US-10-310-914A-223914	Sequence 223914,

ALIGNMENTS

RESULT 1

US-11-112-908-22
; Sequence 22, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 22
; LENGTH: 172147
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-22

Query Match 100.0%; Score 16; DB 7; Length 172147;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGACCCAGGCCACAT 16

Db 116659 AGACCCAGGCCACAT 116674

RESULT 2

US-11-112-908-23
; Sequence 23, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908

```
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23
; LENGTH: 188682
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-23

Query Match      100.0%; Score 16; DB 7; Length 188682;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AGACCCAGGCCCCACAT 16
Db      66280 AGACCCAGGCCCCACAT 66295

RESULT 3
US-10-310-914A-1360219
; Sequence 1360219, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
; APPLICANT: Bentwich, Isaac
; APPLICANT: Shiler, Kvazat
; TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
; FILE REFERENCE: 06087.0200.CPUS01
; CURRENT APPLICATION NUMBER: US/10/310,914A
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 1388402
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1360219
; LENGTH: 22
; TYPE: RNA
; ORGANISM: Human
US-10-310-914A-1360219

Query Match      93.8%; Score 15; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AGACCCAGGCCCCACA 15
Db      8 AGACCCAGGCCCCACA 22

RESULT 4
US-10-310-914A-1360184
; Sequence 1360184, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
; APPLICANT: Bentwich, Isaac
; APPLICANT: Shiler, Kvazat
; TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
; FILE REFERENCE: 06087.0200.CPUS01
; CURRENT APPLICATION NUMBER: US/10/310,914A
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 1388402
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1360184
; LENGTH: 25
; TYPE: RNA
; ORGANISM: Human
```

```
US-10-310-914A-1360184

Query Match      93.8%; Score 15; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AGACCCAGGCCCCACA 15
Db      2 AGACCCAGGCCCCACA 16

RESULT 5
US-10-995-561-13421
; Sequence 13421, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13421
; LENGTH: 403278
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(403278)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-995-561-13421

Query Match      93.8%; Score 15; DB 6; Length 403278;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AGACCCAGGCCCCACA 15
Db      197280 AGACCCAGGCCCCACA 197294

RESULT 6
US-11-101-244-97619
; Sequence 97619, Application US/11101244
; Publication No. US20050246794A1
; GENERAL INFORMATION:
; APPLICANT: Dharmason, Inc.
; APPLICANT: Khvorova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scaringe, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCE: 13499US
; CURRENT APPLICATION NUMBER: US/11/101,244
; CURRENT FILING DATE: 2005-04-07
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary
; SEQ ID NO 97619
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-101-244-97619

Query Match      90.0%; Score 14.4; DB 8; Length 19;
Best Local Similarity 87.5%; Pred. No. 4e+02;
```


GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 16:32:47 ; Search time 58.24 Seconds
(without alignments)
488.341 Million cell updates/sec

Title: US-10-754-446-7

Perfect score: 16

Sequence: 1 agaccaggccacat 16

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgm2_6/ptodata/1/ina/1 COMB.seq.*
2: /cgm2_6/ptodata/1/ina/5 COMB.seq.*
3: /cgm2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgm2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgm2_6/ptodata/1/ina/H COMB.seq.*
6: /cgm2_6/ptodata/1/ina/PCRUS COMB.seq.*
7: /cgm2_6/ptodata/1/ina/PP COMB.seq.*
8: /cgm2_6/ptodata/1/ina/RE COMB.seq.*
9: /cgm2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	100.0	601	3	US-09-949-016-59085
2	16	100.0	2051	3	US-09-949-016-1712
3	16	100.0	15353	3	US-09-949-016-13454
4	15	93.8	2004	3	US-10-104-047-1222
5	15	93.8	56963	3	US-09-949-016-12966
6	15	93.8	56968	3	US-09-949-016-11888
7	15	93.8	254366	3	US-09-822-871-3
8	14.4	90.0	256	3	US-09-513-999C-25589
9	14.4	90.0	533	3	US-09-270-767-9339
10	14.4	90.0	533	3	US-09-270-767-24621
11	14.4	90.0	601	3	US-09-949-016-31585
12	14.4	90.0	601	3	US-09-949-016-31586
13	14.4	90.0	601	3	US-09-949-016-91141
14	14.4	90.0	601	3	US-09-949-016-94698
15	14.4	90.0	601	3	US-09-949-016-94699
16	14.4	90.0	601	3	US-09-949-016-151080
17	14.4	90.0	601	3	US-09-949-016-151081
18	14.4	90.0	1224	3	US-09-581-105-7
19	14.4	90.0	2502	2	US-08-073-384C-7
20	14.4	90.0	2502	2	US-08-254-359A-7
21	14.4	90.0	2502	2	US-08-483-043-7
22	14.4	90.0	2502	2	US-08-481-238-7
23	14.4	90.0	2502	2	US-08-471-066B-7
24	14.4	90.0	2502	2	US-08-484-956-7

25	14.4	90.0	2502	2	US-08-757-653-7	Sequence 7, Appli
26	14.4	90.0	2502	2	US-08-599-491-7	Sequence 7, Appli
27	14.4	90.0	2502	2	US-08-756-386-7	Sequence 7, Appli
28	14.4	90.0	2502	2	US-08-823-516-7	Sequence 7, Appli
29	14.4	90.0	2502	3	US-08-682-853A-7	Sequence 7, Appli
30	14.4	90.0	2502	3	US-08-759-038-7	Sequence 7, Appli
31	14.4	90.0	2502	3	US-08-758-314-7	Sequence 7, Appli
32	14.4	90.0	2502	3	US-09-350-309-7	Sequence 7, Appli
33	14.4	90.0	2502	3	US-08-520-946-7	Sequence 7, Appli
34	14.4	90.0	2502	3	US-09-684-938-7	Sequence 7, Appli
35	14.4	90.0	2502	3	US-09-308-825A-7	Sequence 7, Appli
36	14.4	90.0	2502	3	US-09-758-282B-7	Sequence 7, Appli
37	14.4	90.0	2502	3	US-09-655-378A-7	Sequence 7, Appli
38	14.4	90.0	2502	3	US-09-940-244-7	Sequence 7, Appli
39	14.4	90.0	2502	3	US-09-333-145-7	Sequence 7, Appli
40	14.4	90.0	2502	3	US-09-577-304A-7	Sequence 7, Appli
41	14.4	90.0	2502	3	US-09-381-212-7	Sequence 7, Appli
42	14.4	90.0	2502	3	US-10-081-806-7	Sequence 7, Appli
43	14.4	90.0	2502	3	US-09-713-601A-7	Sequence 7, Appli
44	14.4	90.0	2601	3	US-10-104-047-1515	Sequence 1515, Ap
45	14.4	90.0	3496	3	US-08-660-451A-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1

US-09-949-016-59085
; Sequence 59085, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 59085
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-59085

Query Match 100.0%; Score 16; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGACCCAGGCCACAT 16
|||
Db 449 AGACCCAGGCCACAT 464

RESULT 2

US-09-949-016-1712
; Sequence 1712, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1712
; LENGTH: 2051
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1712

Query Match 100.0%; Score 16; DB 3; Length 2051;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGACCCAGGCCCAT 16
Db 952 AGACCCAGGCCCAT 967

RESULT 3

US-09-949-016-13454
; Sequence 13454, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13454
; LENGTH: 15353
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13454

Query Match 100.0%; Score 16; DB 3; Length 15353;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGACCCAGGCCCAT 16
Db 7582 AGACCCAGGCCCAT 7597

RESULT 4

US-10-104-047-1222
; Sequence 1222, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:

; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1222
; LENGTH: 2004
; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-104-047-1222

Query Match 93.8%; Score 15; DB 3; Length 2004;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GACCCAGGCCCAT 16
Db 1721 GACCCAGGCCCAT 1735

RESULT 5

US-09-949-016-12966/c
; Sequence 12966, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12966
; LENGTH: 56963
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12966

Query Match 93.8%; Score 15; DB 3; Length 56963;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGACCCAGGCCCAT 15
Db 45193 AGACCCAGGCCCAT 45179

RESULT 6

US-09-949-016-11888/c
; Sequence 11888, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11888
; LENGTH: 56968
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-11888

Query Match 93.8%; Score 15; DB 3; Length 56968;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 15:49:05 ; Search time 785.792 Seconds
(without alignments)
1157.425 Million cell updates/sec

Title: US-10-754-446-5

Perfect score: 16

Sequence: 1 tctgcccacagtacct 16

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_on:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_str:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	100.0	790	8 AF305572S3	AF305574 Homo sapi
2	16	100.0	2551	5 AJ720988	AJ720988 Gallus ga
3	16	100.0	7939	4 AF292105	AF292105 Sus scrof
4	16	100.0	13270	8 AF287270	AF287270 Homo sapi
5	16	100.0	45941	8 AL583835	AL583835 Human DNA
6	16	100.0	104062	5 CR855371	CR855371 Chicken D
7	16	100.0	138616	14 AC115525	AC115525 Rattus no
8	16	100.0	155645	14 AC021153	AC021153 Homo sapi
9	16	100.0	166504	8 AC008275	AC008275 Homo sapi
10	16	100.0	173126	8 AC008878	AC008878 Homo sapi
11	16	100.0	188149	14 AC151428	AC151428 Ocolemur
12	16	100.0	190284	14 AC115524	AC115524 Rattus no
13	16	100.0	215397	8 AL355497	AL355497 Human DNA
14	16	100.0	215470	14 AC147936	AC147936 Ocolemur
15	16	100.0	226981	14 AC127822	AC127822 Rattus no
16	16	100.0	256594	14 AC133700	AC133700 Rattus no
17	15	93.8	310	6 CQ687156	CQ687156 Sequence
18	15	93.8	523	10 BV427581	BV427581 S237P6504

C 19	15	93.8	622	10	BV315511	BV315511 S236P6394
C 20	15	93.8	670	10	BV025162	BV025162 S212P6124
C 21	15	93.8	1525	5	CR386310	CR386310 Gallus ga
C 22	15	93.8	4992	5	AJ720888	AJ720888 Gallus ga
C 23	15	93.8	5210	9	AK122332	AK122332 Mus muscu
C 24	15	93.8	6331	13	EMVRPVP	J04374 Eggplant mo
C 25	15	93.8	62000	8	AP001984	AP001984 Homo sapi
C 26	15	93.8	110000	1	CP000082	Continuation (25 o
C 27	15	93.8	110000	14	AC118875	Continuation (3 of
C 28	15	93.8	118103	14	AC009817	AC009817 Homo sapi
C 29	15	93.8	134697	8	AC114738	AC114738 Homo sapi
C 30	15	93.8	138317	8	AP003481	AP003481 Homo sapi
C 31	15	93.8	139512	8	HSJ1169J3	AL049652 Human DNA
C 32	15	93.8	140968	9	BX537299	BX537299 Mouse DNA
C 33	15	93.8	151698	14	AC132644	AC132644 Rattus no
C 34	15	93.8	152816	14	AC079739	AC079739 Homo sapi
C 35	15	93.8	156808	14	AC012407	AC012407 Homo sapi
C 36	15	93.8	161100	14	AC025659	AC025659 Homo sapi
C 37	15	93.8	161179	14	AC083954	AC083954 Homo sapi
C 38	15	93.8	161362	14	AL355503	AL355503 Homo sapi
C 39	15	93.8	167322	14	AC016788	AC016788 Homo sapi
C 40	15	93.8	168789	9	AC127318	AC127318 Mus muscu
C 41	15	93.8	171141	9	AC157651	AC157651 Mus muscu
C 42	15	93.8	172160	9	AC119996	AC119996 Mus muscu
C 43	15	93.8	172772	9	AC121923	AC121923 Mus muscu
C 44	15	93.8	173805	9	AC124548	AC124548 Mus muscu
C 45	15	93.8	175048	14	AC161042	AC161042 Mus muscu

ALIGNMENTS

RESULT 1	AF305572S3	Homo sapiens mucolipin 1 (MCOLN1) gene, exons 3 and 4.	790 bp	DNA	linear	PRI 26-DEC-2000
LOCUS	AF305572S3					
DEFINITION	Homo sapiens mucolipin 1 (MCOLN1) gene, exons 3 and 4.					
ACCESSION	AF305574					
VERSION	AF305574.1	GI:11991200				
KEYWORDS	3 of 8					
SEGMENT	Homo sapiens (human)					
SOURCE	Homo sapiens					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 790)					
AUTHORS	Bargal,R., Avidan,N., Ben-Asher,E., Olender,Z., Zeigler,M., Frumkin,A., Raas-Rothschild,A., Glueman,G., Lancet,D. and Bach,G.					
TITLE	Identification of the gene causing mucopolidiosis type IV					
JOURNAL	Nat. Genet. 26 (1), 118-123 (2000)					
PUBMED	10973263					
REFERENCE	2 (bases 1 to 790)					
AUTHORS	Bargal,R., Avidan,N., Ben-Asher,E., Olender,A., Zeigler,M., Frumkin,A., Raas-Rothschild,A., Glueman,G., Lancet,D. and Bach,G.					
TITLE	Direct Submission					
JOURNAL	Submitted (13-SEP-2000) Molecular Genetics, The Weizmann Institute of Science, P. O. Box 26, Rehovot 76100, Israel					
FEATURES	Location/Qualifiers					
source	1..790					
exon	/organism="Homo sapiens"					
	/mol_type="genomic DNA"					
	/db_xref="taxon:9606"					
	267..434					
	/gene="MCOLN1"					
	/number=3					
exon	589..757					
	/gene="MCOLN1"					
	/number=4					
ORIGIN						
Query Match	100.0%;	Score 16;	DB 8;	Length 790;		
Best Local Similarity	100.0%;	Pred. No. 1.6e+02;	Mismatches 0;	Indels 0;	Gaps 0;	
Matches 16;	Conservative 0;					

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 15:55:50 ; Search time 1789.31 Seconds
(without alignments)
418.369 Million cell updates/sec

Title: US-10-754-446-5
Perfect score: 16
Sequence: 1 tctgccacagatcact 16

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_hc.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_est7.*
9: gb_gss1.*
10: gb_gss2.*
11: gb_gss3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	16	100.0	154	9 BH271054	BH271054 CH230-162
2	16	100.0	480	2 BB856429	BB856429 BB856429
3	16	100.0	525	9 AQ669413	AQ669413 HS_5407_B
C 4	16	100.0	696	1 AJ727930	AJ727930 AJ727930
5	16	100.0	745	1 AJ727936	AJ727936 AJ727936
6	16	100.0	762	5 BU306972	BU306972 603612510
7	16	100.0	774	8 DR431852	DR431852 nax48ff12
8	16	100.0	806	5 BU383135	BU383135 603961896
9	16	100.0	1008	5 BU420692	BU420692 603231561
C 10	15	93.8	224	8 CV972084	CV972084 LRRGE0208
C 11	15	93.8	269	8 CV972398	CV972398 LRRGE0239
12	15	93.8	279	1 BB007743	BB007743 BB007743
13	15	93.8	279	8 CV972703	CV972703 LRRGE0270
14	15	93.8	281	8 CV888115	CV888115 LRRGE0000
C 15	15	93.8	361	8 CV972086	CV972086 LRRGE0208
C 16	15	93.8	388	5 BY619013	BY619013 BY619013
C 17	15	93.8	420	9 BH273644	BH273644 CH230-70E
C 18	15	93.8	436	8 CV888113	CV888113 LRRGE0000
19	15	93.8	475	8 CV888114	CV888114 LRRGE0000
C 20	15	93.8	492	9 A2822326	A2822326 2M0095J01
C 21	15	93.8	501	8 BH117609	BH117609 RPCI-24-2
22	15	93.8	507	9 A2768935	A2768935 1M0569J09

```

23 15 93.8 517 9 AZ585292
24 15 93.8 525 9 AQ357179
c 25 15 93.8 591 10 CE537462
26 15 93.8 595 2 BG404774
c 27 15 93.8 599 1 AI412276
28 15 93.8 625 7 CK620962
29 15 93.8 672 10 CE809412
30 15 93.8 714 5 BY764892
31 15 93.8 727 5 BU724937
c 32 15 93.8 727 5 BU139411
c 33 15 93.8 732 1 AJ724931
c 34 15 93.8 734 10 CE812263
c 35 15 93.8 745 11 CR081738
c 36 15 93.8 753 10 AG457089
c 37 15 93.8 763 11 CR228261
c 38 15 93.8 784 3 BI331708
c 39 15 93.8 784 7 CF924020
40 15 93.8 790 7 CN234992
41 15 93.8 797 7 CK315394
42 15 93.8 870 11 CR118310
c 43 15 93.8 999 2 BF976532
c 44 15 93.8 1001 5 BU192451
45 15 93.8 1184 7 CK167560

```

ALIGNMENTS

```

RESULT 1
BH271054/c 154 bp DNA linear GSS 30-NOV-2001
LOCUS BH271054
DEFINITION CH230-162L17.TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone
CH230-162L17, genomic survey sequence.
ACCESSION BH271054
VERSION BH271054.1 GI:17183456
KEYWORDS GSS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Murinae; Rattus.
1 (bases 1 to 154)
Zhao,S., Shetty,J., Shatsman,S., Tsengay,G., Geer,K.,
Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D.,
Riggs,F., de Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 EcoRI segment
Unpublished (1999)
Other_GSSs: CH230-162L17.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering_information.htm). BAC end
page: http://www.tigr.org/tadb/bac_ends/rat/bac_end_intro.html
Plate: 162 row: L column: 17
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1..154
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SsNHd/MCW"
/db_xref="taxon:10116"
/clone="CH230-162L17"
/sex="Female"
/cell_type="Brain"

```

/clone_lib="CHORI-230 Segment 1"
 /note="vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
 CHORI-230 Rat (BN/SeNHsd/MCW) BAC library produced by
 Pieter de Jong"

ORIGIN

Query Match 100.0%; Score 16; DB 9; Length 154;
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGCCCACAGTACCT 16
 |||||
 Db 150 TCTGCCCACAGTACCT 135

RESULT 2

BB856429

LOCUS

BB856429 BB856429 480 bp mRNA linear EST 26-NOV-2001
 cDNA clone G370034110 5', mRNA sequence.

ACCESSION

BB856429

VERSION

BB856429.1

KEYWORDS

EST.

SOURCE

Mus musculus

ORGANISM

Mus musculus

REFERENCE

AUTHORS

1 (bases 1 to 480)
 Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
 Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K.,
 Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
 Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,
 Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,
 Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
 Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
 Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A.,
 Muramatsu,M. and Hayashizaki,Y.
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
 2001)

TITLE

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
 2001)

JOURNAL

COMMENT

Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216

Email: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
 Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
 Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
 Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
 and Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multipipillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
 Sugahara,Y. and Hayashizaki,Y.

Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Please visit our web site (http://genome.gsc.riken.go.jp) for
 further details.

e mouse tissues.

FEATURES

source

1. .480
 /organism="Mus musculus"
 /mol_type="mRNA"

/strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="G370034110"
 /cell_type="B16 F10Y cells"
 /clone_lib="RIKEN full-length enriched, B16 F10Y cells"

ORIGIN

Query Match 100.0%; Score 16; DB 2; Length 480;
 Best Local Similarity 100.0%; Pred. No. 1.6e+03;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGCCCACAGTACCT 16
 |||||
 Db 366 TCTGCCCACAGTACCT 381

RESULT 3

AQ669413

LOCUS

AQ669413 HS 5407.B1.F10.T7A.RPCI-11 Human Male BAC Library Homo sapiens
 genomic clone Plate=983 Col=19 Row=L, genomic survey sequence.

ACCESSION

AQ669413

VERSION

AQ669413.1

KEYWORDS

GSS.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

1 (bases 1 to 525)
 Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
 Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
 Hood,L.
 Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 10449764

JOURNAL

COMMENT

Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
 or from Research Genetics (info@resgen.com). BAC end Web Server:
 http://www.htsc.washington.edu
 Plate: 983 row: L column: 19
 Seq primer: T7
 Class: BAC ends
 High quality sequence stop: 525.

FEATURES

source

1. .525
 /organism="Homo sapiens"
 /mol_type="genomic DNA"

/db_xref="taxon:9606"
 /clone="Plate=983 Col=19 Row=L"

/sex="male"

/clone_lib="RPCI-11 Human Male BAC Library"

/note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;
 Male blood DNA was isolated from one randomly chosen donor
 and partially digested with a combination of EcoRI and
 EcoRI Methylase. Size selected DNA was cloned into the
 pBACe3.6 vector at EcoRI sites"

ORIGIN

Query Match 100.0%; Score 16; DB 9; Length 525;
 Best Local Similarity 100.0%; Pred. No. 1.6e+03;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 14:15:14 ; Search time 217.088 Seconds
(without alignments)
491.207 Million cell updates/sec

Title: US-10-754-446-5

Perfect score: 16

Sequence: 1 tctgcccacagtaacct 16

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001as.*
5: Geneseq2001bs.*
6: Geneseq2002as.*
7: Geneseq2002bs.*
8: Geneseq2003as.*
9: Geneseq2003bs.*
10: Geneseq2003cs.*
11: Geneseq2003ds.*
12: Geneseq2004as.*
13: Geneseq2004bs.*
14: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16	100.0	16	14	AEB28961 Human MCO
2	16	100.0	11365	4	Aak73827 Human imm
3	16	100.0	13270	9	Adb84283 Human muc
4	16	100.0	20046	4	Aak73826 Human imm
5	15	93.8	177531	8	Acf62732 Cancer ba
6	15	93.8	177531	8	Adb20847 MRPI base
7	15	93.8	177531	10	Adb87936 Human UGT
8	15	93.8	177531	10	Adb96919 Human MDR
9	15	93.8	177531	10	Adb92110 Human MDR
10	15	93.8	177531	10	Adh74617 Human BAC
11	14.4	90.0	405	4	Aal35222 Human mus
12	14.4	90.0	405	8	Abx58210 cDNA enco
13	14.4	90.0	405	12	Adj27937 Human end
c 14	14.4	90.0	465	9	Ach32754 Human end
c 15	14.4	90.0	501	10	Adb49981 Primary r
c 16	14.4	90.0	548	13	Aq50585 Novel can
c 17	14.4	90.0	570	4	Aal13764 Human bre
c 18	14.4	90.0	789	2	Av65231 DNA seque
c 19	14.4	90.0	835	6	Ab190105 Human pol

c	20	14.4	90.0	897	11	ACN83833	Acn83833 Breast ca
	21	14.4	90.0	974	6	AAD31825	Aad31825 Human pan
	22	14.4	90.0	1044	13	ADS56881	Ads56881 Bacterial
c	23	14.4	90.0	1692	4	AAK79833	Aak79833 Human imm
	24	14.4	90.0	1849	5	AAS83682	Aas83682 DNA enco
	25	14.4	90.0	2214	13	ADU01699	Adu01699 Novel hum
	26	14.4	90.0	2228	6	ABQ99407	Abq99407 Human cod
	27	14.4	90.0	2297	12	ADQ59977	Adq59977 T cell ac
	28	14.4	90.0	2317	6	ABS76414	Abs76414 cDNA enco
	29	14.4	90.0	2317	6	ABV99386	Abv99386 Human NOV
	30	14.4	90.0	2317	13	ADU06214	Adu06214 Novel bro
	31	14.4	90.0	2380	4	AAL36812	Aal36812 Human mus
	32	14.4	90.0	2380	8	ABX59800	Abx59800 cDNA enco
	33	14.4	90.0	2380	12	ADJ30550	Adj30550 Human mus
	34	14.4	90.0	2394	4	AAL36811	Aal36811 Human mus
	35	14.4	90.0	2394	8	ABX59799	Abx59799 cDNA enco
	36	14.4	90.0	2394	12	ADJ30549	Adj30549 Human mus
	37	14.4	90.0	2452	6	AAI71445	Aai71445 TNFR/NGFR
	38	14.4	90.0	2569	10	ADB62298	Adb62298 Human cDN
	39	14.4	90.0	2788	4	AAH17843	Aah17843 Human cDN
	40	14.4	90.0	2816	13	ACN38987	Acn38987 Tumour-as
	41	14.4	90.0	3025	2	AAQ13337	Aaq13337 D1 dopami
	42	14.4	90.0	3025	2	AAT63657	Aat63657 D1 dopami
c	43	14.4	90.0	3778	6	ABZ11278	Abz11278 Human pol
c	44	14.4	90.0	3778	12	ADM43796	Adm43796 Novel hum
c	45	14.4	90.0	4376	10	ADF50290	Adf50290 Human PFM

ALIGNMENTS

RESULT 1

AEB28961
ID AEB28961 standard; DNA; 16 BP.
XX
AC AEB28961;
XX
DT 22-SEP-2005 (first entry)
XX
DE Human MCOLN1 gene probe IVS WT SEQ ID NO:5.
XX
KW mucolipidosis IV; high throughput screening; mucolipin-1; MCOLN1; probe;
KW ss.
XX
OS Homo sapiens.
PN US2005153300-A1.
XX
PD 14-JUL-2005.
XX
PF 09-JAN-2004; 2004US-00754446.
XX
PR 09-JAN-2004; 2004US-00754446.
XX
XX (QUES-) QUEST DIAGNOSTICS INC.
XX
XX Sun W, Hantash F;
XX
XX WPI; 2005-521160/53.

Diagnosing mucolipidosis (ML) IV mutant sequence in nucleic acids, by amplifying the nucleic acid, detecting amplified product with labeled oligonucleotide probes via a change in fluorescence which indicates the presence of an ML IV mutant.

Claim 3; SEQ ID NO 5; 15pp; English.

The invention relates to a method (M1) for detecting the presence of mucolipidosis IV (ML IV) mutant sequences in a nucleic acid. (M1) involves contacting the nucleic acid with oligonucleotide primers and probes, conducting amplification by temperature cycling and monitoring the accumulation of amplified nucleic acid by detecting an increase in donor fluorophore fluorescence or decrease in acceptor fluorophore

CC fluorescence, which indicates the presence of ML IV mutant sequence. Also
 CC described: (1) an oligonucleotide (ON1) comprising a sequence
 CC complementary to the coding or non-coding strand of the mucolin-1
 CC (MCOLN1) gene; and (2) a kit (K1) for amplifying sequences of ML IV
 CC sequences, comprising ON1. (M1) is useful for determining the presence of
 CC ML IV mutant sequence in a nucleic acid, and/or wild-type sequences. (M1)
 CC is useful for genetic counseling of individuals at risk for ML IV or at
 CC risk to having offspring afflicted with ML IV. The MCOLN1 gene is located
 CC on chromosome 19p13. The present sequence represents a probe for the
 CC MCOLN1 gene, which is used in the exemplification of the present
 CC invention.

XX
 SQ Sequence 16 BP; 3 A; 7 C; 2 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 16; DB 14; Length 16;

Best Local Similarity 100.0%; Pred. No. 1.3e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGCCACAGTACCT 16

Db 1 TCTGCCACAGTACCT 16

RESULT 2

AAK73827

ID AAK73827 standard; DNA; 11365 BP.

XX AC AAK73827;

XX DT 07-NOV-2001 (first entry)

XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:28639.

XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

XX KW Cytostatic; gene therapy; vaccine; metastasis; ds.

XX OS Homo sapiens.

XX PN WC200157182-A2.

XX XX 09-AUG-2001.

XX XX 17-JAN-2001; 2001WO-US001354.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225575P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239335P.
 PR 13-OCT-2000; 2000US-0239337P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 16:41:33 ; Search time 386.432 Seconds
(without alignments)
342.389 Million cell updates/sec

Title: US-10-754-446-5

Perfect score: 16

Sequence: 1 tctgcccacagctacct 16

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main:
1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubna/US10D_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubna/US10E_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	100.0	16	9	US-10-754-446-5
2	16	100.0	653	4	US-09-925-065A-696499
3	16	100.0	653	4	US-09-925-065A-696500
4	16	100.0	13270	3	US-09-851-494B-1
5	15.6	97.5	600	9	US-10-972-079-69742
6	15.6	97.5	600	9	US-10-972-079-69743
7	15.6	97.5	600	9	US-10-972-079-69744
8	15.6	97.5	600	9	US-10-972-079-69745
9	15	93.8	310	7	US-10-242-535A-32082
10	15	93.8	310	7	US-10-085-783A-32082
11	15	93.8	533	5	US-10-027-632-224114
12	15	93.8	533	6	US-10-027-632-224114
13	15	93.8	177531	8	US-10-484-597-660
14	14.4	90.0	25	7	US-10-719-956-630855
15	14.4	90.0	275	9	US-10-756-149-4182
16	14.4	90.0	295	8	US-10-674-144A-10971
17	14.4	90.0	297	3	US-09-783-590-8724
18	14.4	90.0	312	6	US-10-062-674-400
19	14.4	90.0	388	3	US-09-783-590-8696
20	14.4	90.0	390	3	US-09-783-590-8711
21	14.4	90.0	390	3	US-09-783-590-8711
22	14.4	90.0	396	3	US-09-783-590-8813
23	14.4	90.0	396	3	US-09-783-590-8589

24	14.4	90.0	405	3	US-09-764-877-564	Sequence 564, App
25	14.4	90.0	405	6	US-10-242-515-564	Sequence 564, App
26	14.4	90.0	411	4	US-09-925-065A-54604	Sequence 54604, A
27	14.4	90.0	452	4	US-09-925-065A-256794	Sequence 256794, A
28	14.4	90.0	452	4	US-09-925-065A-256795	Sequence 256795, A
29	14.4	90.0	452	4	US-09-925-065A-256796	Sequence 256796, A
30	14.4	90.0	452	3	US-09-918-995-19966	Sequence 19966, A
31	14.4	90.0	533	4	US-09-925-065A-112177	Sequence 112177, A
32	14.4	90.0	563	4	US-09-925-065A-197600	Sequence 197600, A
33	14.4	90.0	563	4	US-09-925-065A-197601	Sequence 197601, A
34	14.4	90.0	585	5	US-10-027-632-229737	Sequence 229737, A
35	14.4	90.0	585	5	US-10-027-632-229738	Sequence 229738, A
36	14.4	90.0	585	6	US-10-027-632-229737	Sequence 229737, A
37	14.4	90.0	585	6	US-10-027-632-229738	Sequence 229738, A
38	14.4	90.0	600	9	US-10-972-079-40815	Sequence 40815, A
39	14.4	90.0	606	4	US-09-925-065A-495309	Sequence 495309, A
40	14.4	90.0	622	4	US-09-925-065A-402883	Sequence 402883, A
41	14.4	90.0	626	4	US-09-925-065A-539208	Sequence 539208, A
42	14.4	90.0	626	4	US-09-925-065A-539209	Sequence 539209, A
43	14.4	90.0	630	5	US-10-027-632-275537	Sequence 275537, A
44	14.4	90.0	630	6	US-10-027-632-275537	Sequence 275537, A
45	14.4	90.0	655	4	US-09-925-065A-397489	Sequence 397489, A

ALIGNMENTS

RESULT 1
US-10-754-446-5
; Sequence 5, Application US/10754446
; Publication No. US20050153300A1
; GENERAL INFORMATION:
; APPLICANT: SUN, WEIMIN
; APPLICANT: HANTASH, PERAS
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DETECTION OF
; TITLE OF INVENTION: MUCOLIPIDOSIS IV MUTATIONS
; FILE REFERENCE: 034827-2301
; CURRENT APPLICATION NUMBER: US/10754.446
; CURRENT FILING DATE: 2004-01-09
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 3.2
; SEQ ID NO 5
; LENGTH: 16
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Probe
US-10-754-446-5

Query Match 100.0%; Score 16; DB 9; Length 16;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGCCCCACAGTACCT 16
| | | | | | | | | | | | | | | |
Db 1 TCTGCCCCACAGTACCT 16

RESULT 2
US-09-925-065A-696499
; Sequence 696499, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 168827.135
; CURRENT APPLICATION NUMBER: US/09/925.065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20

; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 696499
; LENGTH: 653
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-696499

Query Match 100.0%; Score 16; DB 4; Length 653;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGCCCCACAGTACCT 16
|||||
Db 137 TCTGCCCCACAGTACCT 152

RESULT 3

US-09-925-065A-696500
; Sequence 696500, Application US/09925065A
; Publication No. US20030228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 696500
; LENGTH: 653
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-696500

Query Match 100.0%; Score 16; DB 4; Length 653;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGCCCCACAGTACCT 16
|||||
Db 137 TCTGCCCCACAGTACCT 152

RESULT 4

US-09-851-494B-1
; Sequence 1, Application US/09851494B
; Publication No. US20030064363A1
; GENERAL INFORMATION:
; APPLICANT: ML4 Foundation
; APPLICANT: Goldin, Ehud
; APPLICANT: Slaugenhaupt, Susan A.
; APPLICANT: Sun, Mei
; APPLICANT: Acierno, James S.
; TITLE OF INVENTION: A Gene Encoding A New TRP Channel is Mutated in Mucopolipidosis IV
; FILE REFERENCE: 3394/1H57U51

; CURRENT APPLICATION NUMBER: US/09/851,494B
; CURRENT FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 13270
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-851-494B-1

Query Match 100.0%; Score 16; DB 3; Length 13270;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGCCCCACAGTACCT 16
|||||
Db 5525 TCTGCCCCACAGTACCT 5540

RESULT 5

US-10-972-079-69742
; Sequence 69742, Application US/10972079
; Publication No. US20050153317A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: ROSENFELD, David
; APPLICANT: KERR, Richard
; APPLICANT: BATES, Stephen
; APPLICANT: HOLM, Tom
; TITLE OF INVENTION: METHODS & SYSTEMS FOR INFERRING TRAITS TO BREED & MANAGE NON-BEEP
; FILE OF INVENTION: LIVESTOCK
; FILE REFERENCE: MM1110-2
; CURRENT APPLICATION NUMBER: US/10/972,079
; CURRENT FILING DATE: 2004-10-22
; PRIOR APPLICATION NUMBER: US 60/514,333
; PRIOR FILING DATE: 2003-10-24
; NUMBER OF SEQ ID NOS: 96631
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69742
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Chicken 19866994342124_1
US-10-972-079-69742

Query Match 97.5%; Score 15.6; DB 9; Length 600;
Best Local Similarity 93.8%; Pred. No. 1.2e+02;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGCCCCACAGTACCT 16
:|||||
Db 461 WCTGCCCCACAGTACCT 476

RESULT 6

US-10-972-079-69743
; Sequence 69743, Application US/10972079
; Publication No. US20050153317A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: ROSENFELD, David
; APPLICANT: KERR, Richard
; APPLICANT: BATES, Stephen
; APPLICANT: HOLM, Tom
; TITLE OF INVENTION: METHODS & SYSTEMS FOR INFERRING TRAITS TO BREED & MANAGE NON-BEEP
; FILE OF INVENTION: LIVESTOCK
; FILE REFERENCE: MM1110-2
; CURRENT APPLICATION NUMBER: US/10/972,079
; CURRENT FILING DATE: 2004-10-22
; PRIOR APPLICATION NUMBER: US 60/514,333
; PRIOR FILING DATE: 2003-10-24
; NUMBER OF SEQ ID NOS: 96631

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 16:59:57 ; Search time 196.864 Seconds
(without alignments) 65.779 Million cell updates/sec

Title: US-10-754-446-5

Perfect score: 16

Sequence: 1 tctgccacagtacct 16

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6038814 seqs, 404674181 residues

Total number of hits satisfying chosen parameters: 12077628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

```
Database : Published Applications NA New:*
1: /cgn2_6/ptodata/2/pubpna/US08 NEW PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06 NEW PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07 NEW PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US09 NEW PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US10 NEW PUB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US11 NEW PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US11 NEW PUB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US11 NEW PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US60 NEW PUB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	14.4	90.0	69	6	US-10-310-914A-10301	Sequence 10301, A
2	14.4	90.0	69	6	US-10-310-914A-16743	Sequence 16743, A
C 3	14.4	90.0	1126	6	US-10-750-185-61765	Sequence 61765, A
C 4	14.4	90.0	1126	6	US-10-750-623-61765	Sequence 61765, A
C 5	14.4	90.0	2095	6	US-10-750-185-48943	Sequence 48943, A
C 6	14.4	90.0	2095	6	US-10-750-623-48943	Sequence 48943, A
C 7	14.4	90.0	3396	6	US-10-750-185-57338	Sequence 57338, A
8	14.4	90.0	3396	6	US-10-750-623-57338	Sequence 57338, A
C 9	14.4	90.0	40644	6	US-10-995-561-13480	Sequence 13480, A
10	14.4	90.0	340000	7	US-11-102-978-3	Sequence 3, Appl
11	14	87.5	201	6	US-10-995-561-18449	Sequence 18449, A
C 12	14	87.5	201	6	US-10-995-561-24263	Sequence 24263, A
C 13	14	87.5	201	6	US-10-995-561-24481	Sequence 24481, A
14	14	87.5	600	7	US-11-136-527-5051	Sequence 5051, Ap
C 15	14	87.5	624	7	US-11-136-527-955	Sequence 955, App
C 16	14	87.5	625	6	US-10-750-185-54379	Sequence 54379, A
C 17	14	87.5	625	6	US-10-750-623-54379	Sequence 54379, A
C 18	14	87.5	159138	6	US-10-995-561-13230	Sequence 13230, A
C 19	14	87.5	241805	6	US-10-995-561-13215	Sequence 13215, A
C 20	13.4	83.8	25	7	US-11-121-849-620404	Sequence 620404, A
C 21	13.4	83.8	25	7	US-11-121-849-638454	Sequence 638454, A
C 22	13.4	83.8	201	6	US-10-995-561-18604	Sequence 18604, A
C 23	13.4	83.8	447	7	US-11-136-527-3908	Sequence 3908, Ap

ALIGNMENTS

RESULT 1

```

US-10-310-914A-10301
; Sequence 10301, Application US/10310914A
; Publication No. US2006000322A1
; GENERAL INFORMATION:
; APPLICANT: Bentwich, Isaac
; APPLICANT: Shiller, Kvuzat
; TITLE OF INVENTION: Bioinformatically detectable
; TITLE OF INVENTION: uses thereof
; FILE REFERENCE: 06087.0200.CPUS01
; CURRENT APPLICATION NUMBER: US/10/310,914A
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 1388402
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 10301
; LENGTH: 69
; TYPE: RNA
; ORGANISM: Human
US-10-310-914A-10301

```

Query Match 90.0%; Score 14.4; DB 6; Length 69;
Best Local Similarity 68.8%; Pred. No. 73;
Matches 11: Conservative 4; Mismatches 1; Indels

QY 1 TCTGCCCACAGTACCT 16
:|:| | | | | | | :
Db 2A UCTGCACACAGUACCT 43

RESULT 2

```

US-10-310-914A-16743
; Sequence 16743, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
; APPLICANT: Bentwich, Isaac
; APPLICANT: Shiler, Kvuzat
; TITLE OF INVENTION: Bioinformatically detectable
; uses thereof
; FILE REFERENCE: 06087.0200, CPUS01
; CURRENT APPLICATION NUMBER: US/10/310,914A
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 1389402
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 16743
; LENGTH: 69
; TYPE: RNA

```

```
; ORGANISM: Human
US-10-310-914A-16743

Query Match          90.0%; Score 14.4; DB 6; Length 69;
Best Local Similarity 93.8%; Pred. No. 73;
Matches 11; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGCCCCACAGTACCT 16
   :||| |||||:|:|:|
Db 28 UCUGCACACAGUACCU 43

RESULT 3
US-10-750-185-61765/c
; Sequence 61765, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10750,185
; PRIOR FILING DATE: 2003-12-31
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 61765
; LENGTH: 1126
; TYPE: DNA
; ORGANISM: Bovine 19866881607649
US-10-750-185-61765

Query Match          90.0%; Score 14.4; DB 6; Length 1126;
Best Local Similarity 93.8%; Pred. No. 80;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGCCCCACAGTACCT 16
   ||||| ||||| |||||
Db 599 TCTGCCACACAGTACCT 584

RESULT 4
US-10-750-623-61765/c
; Sequence 61765, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10750,623
; PRIOR FILING DATE: 2003-12-31
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 61765
; LENGTH: 1126
; TYPE: DNA
; ORGANISM: Bovine 19866881607649
US-10-750-623-61765

Query Match          90.0%; Score 14.4; DB 6; Length 1126;
Best Local Similarity 93.8%; Pred. No. 81;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGCCCCACAGTACCT 16
   ||||| ||||| |||||
Db 1253 TCTGCCACACAGTACCT 1238

RESULT 5
US-10-750-185-48943/c
; Sequence 48943, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 48943
; LENGTH: 2095
; TYPE: DNA
; ORGANISM: Bovine 19866881228818
US-10-750-185-48943

Query Match          90.0%; Score 14.4; DB 6; Length 2095;
Best Local Similarity 93.8%; Pred. No. 81;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGCCCCACAGTACCT 16
   ||||| ||||| |||||
Db 1253 TCTGCCACACAGTACCT 1238

RESULT 6
US-10-750-623-48943/c
; Sequence 48943, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10750,623
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 48943
; LENGTH: 2095
; TYPE: DNA
; ORGANISM: Bovine 19866881228818
US-10-750-623-48943

Query Match          90.0%; Score 14.4; DB 6; Length 2095;
Best Local Similarity 93.8%; Pred. No. 81;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 16:32:47 ; Search time 58.24 Seconds
(without alignments)
488.341 Million cell updates/sec

Title: US-10-754-446-5

Perfect score: 16

Sequence: 1 tctgccacagctacct 16

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PCUTUS COMB.seq.*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	100.0	15353	3	US-09-949-016-13454
2	16	100.0	110585	3	US-09-949-016-13427
3	14.4	90.0	601	3	US-09-949-016-22226
4	14.4	90.0	601	3	US-09-949-016-22227
5	14.4	90.0	601	3	US-09-949-016-22228
6	14.4	90.0	601	3	US-09-949-016-22229
7	14.4	90.0	601	3	US-09-949-016-22230
8	14.4	90.0	601	3	US-09-949-016-22231
9	14.4	90.0	601	3	US-09-949-016-22232
10	14.4	90.0	601	3	US-09-949-016-22233
11	14.4	90.0	601	3	US-09-949-016-38209
12	14.4	90.0	601	3	US-09-949-016-38210
13	14.4	90.0	601	3	US-09-949-016-38211
14	14.4	90.0	601	3	US-09-949-016-51333
15	14.4	90.0	601	3	US-09-949-016-51334
16	14.4	90.0	601	3	US-09-949-016-51335
17	14.4	90.0	601	3	US-09-949-016-51336
18	14.4	90.0	601	3	US-09-949-016-51337
19	14.4	90.0	601	3	US-09-949-016-51338
20	14.4	90.0	601	3	US-09-949-016-51339
21	14.4	90.0	601	3	US-09-949-016-51340
22	14.4	90.0	601	3	US-09-949-016-68413
23	14.4	90.0	601	3	US-09-949-016-68414
24	14.4	90.0	601	3	US-09-949-016-68415

```

c 25 14.4 90.0 601 3 US-09-949-016-169541 Sequence 169541,
c 26 14.4 90.0 601 3 US-09-949-016-169542 Sequence 169542,
c 27 14.4 90.0 601 3 US-09-949-016-169543 Sequence 169543,
c 28 14.4 90.0 789 3 US-08-986-765-4 Sequence 4, Appli
c 29 14.4 90.0 2569 3 US-10-104-047-452 Sequence 452, App
c 30 14.4 90.0 2883 3 US-09-614-221A-371 Sequence 371, Appli
c 31 14.4 90.0 3025 2 US-08-444-734A-1 Sequence 1, Appli
c 32 14.4 90.0 3778 3 US-09-799-451-160 Sequence 160, App
c 33 14.4 90.0 4376 3 US-10-200-012-15 Sequence 15, Appli
c 34 14.4 90.0 4521 3 US-09-533-494A-18 Sequence 18, Appli
c 35 14.4 90.0 8461 3 US-09-949-016-13428 Sequence 13428, A
c 36 14.4 90.0 12055 3 US-09-949-016-13233 Sequence 13233, A
c 37 14.4 90.0 67479 3 US-09-949-016-11804 Sequence 11804, A
c 38 14.4 90.0 71119 3 US-09-949-016-15358 Sequence 15358, A
c 39 14.4 90.0 77661 3 US-09-949-016-12770 Sequence 12770, A
c 40 14.4 90.0 77663 3 US-09-949-016-13751 Sequence 13751, A
c 41 14.4 90.0 121427 3 US-09-949-016-11950 Sequence 11950, A
c 42 14.4 90.0 121433 3 US-09-949-016-13230 Sequence 13230, A
c 43 14.4 90.0 160759 3 US-09-949-016-16514 Sequence 16514, A
c 44 14 87.5 601 3 US-09-949-016-117072 Sequence 117072,
c 45 14 87.5 601 3 US-09-949-016-152388 Sequence 152388,

```

ALIGNMENTS

RESULT 1

```

US-09-949-016-13454
; Sequence 13454, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13454
; LENGTH: 15353
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13454

```

```

Query Match 100.0%; Score 16; DB 3; Length 15353;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TCTGCCACAGCTACT 16

```

|||||
Db 6125 TCTGCCACAGCTACT 6140

```

RESULT 2

```

US-09-949-016-13427/c
; Sequence 13427, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20

```

; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13427
; LENGTH: 110585
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(110585)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13427

Query Match 100.0%; Score 16; DB 3; Length 110585;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGCCCCACAGTACCT 16
Db 7380 TCTGCCCCACAGTACCT 7365

RESULT 3

US-09-949-016-22226
; Sequence 22226, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 22226
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-22226

Query Match 90.0%; Score 14.4; DB 3; Length 601;
Best Local Similarity 93.8%; Pred. No. 2.4e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGCCCCACAGTACCT 16
Db 77 TCTGCCCCACATACCT 92

RESULT 4

US-09-949-016-22227
; Sequence 22227, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 22227
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-22227

Query Match 90.0%; Score 14.4; DB 3; Length 601;
Best Local Similarity 93.8%; Pred. No. 2.4e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGCCCCACAGTACCT 16
Db 94 TCTGCCCCACAATACCT 109

RESULT 5

US-09-949-016-22228
; Sequence 22228, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 22228
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-22228

Query Match 90.0%; Score 14.4; DB 3; Length 601;
Best Local Similarity 93.8%; Pred. No. 2.4e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGCCCCACAGTACCT 16
Db 464 TCTGCCCCACAATACCT 479

RESULT 6

US-09-949-016-22229
; Sequence 22229, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 15:49:05 ; Search time 1031.35 Seconds
(without alignments)
1157.425 Million cell updates/sec

Title: US-10-754-446-4

Perfect score: 21
Sequence: 1 ctcaccgtgctgaagacact 21

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sts.*
11: gb_av.*
12: gb_un.*
13: gb_vl.*
14: gb_htg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	21	100.0	579	8 AF30557285	AF305576 Homo sapi
C 2	21	100.0	13270	8 AF287270	AF287270 Homo sapi
C 3	21	100.0	155645	14 AC021153	AC021153 Homo sapi
C 4	21	100.0	173126	8 AC008878	AC008878 Homo sapi
C 5	18.4	87.6	1740	6 AX280021	AX280021 Sequence
C 6	18.4	87.6	1829	6 CQ719994	CQ719994 Sequence
C 7	18.4	87.6	2004	8 AF249319	AF249319 Homo sapi
C 8	18.4	87.6	2037	8 HSA293970	HSA293970 Homo sapi
C 9	18.4	87.6	2049	8 AK026102	AK026102 Homo sapi
C 10	18.4	87.6	2051	8 AF287269	AF287269 Homo sapi
C 11	18.4	87.6	2052	6 AX083508	AX083508 Sequence
C 12	18.4	87.6	2063	8 AK222673	AK222673 Homo sapi
C 13	18.4	87.6	2078	8 BC005149	BC005149 Homo sapi
C 14	18.4	87.6	2094	6 BD233734	BD233734 31 human
C 15	18.4	87.6	2095	6 AX280019	AX280019 Sequence
C 16	18.4	87.6	2272	8 HSA293659	HSA293659 Homo sapi
C 17	18.4	87.6	165079	14 AC128011	AC128011 Rattus no
C 18	18.4	87.6	195753	14 AC128091	AC128091 Rattus no

19	18.4	87.6	217148	9 AL773540	AL773540 Mouse DNA
C 20	18.4	87.6	224426	14 AC108247	AC108247 Rattus no
C 21	18.4	87.6	234616	14 AC112952	AC112952 Mus muscu
C 22	18.4	87.6	241828	14 AC106575	AC106575 Rattus no
C 23	18	85.7	10155	1 AE012227	AE012227 Xanthomon
C 24	18	85.7	10446	1 AE011765	AE011765 Xanthomon
C 25	18	85.7	110000	1 AE013598	AE013598 Xanthomon
C 26	18	85.7	110000	1 CP000050	Continuation (20 o
C 27	17.8	84.8	75511	14 AC153278	AC153278 Bos tauru
C 28	17.8	84.8	107113	14 AC142559	AC142559 Takifugu
C 29	17.8	84.8	110000	1 AY596297	Continuation (16 o
C 30	17.8	84.8	145595	14 AC125926	AC125926 Rattus no
C 31	17.8	84.8	169812	14 AC150169	AC150169 Gallus ga
C 32	17.8	84.8	183916	8 AF111169	AF111169 Homo sapi
C 33	17.8	84.8	186330	8 AC007376	AC007376 Homo sapi
C 34	17.8	84.8	198278	14 AC150063	AC150063 Gallus ga
C 35	17.8	84.8	212150	14 AC130048	AC130048 Rattus no
C 36	17.8	84.8	224635	14 AC094848	AC094848 Rattus no
C 37	17.8	84.8	231625	14 AC107597	AC107597 Rattus no
C 38	17.8	84.8	232859	14 AC118408	AC118408 Rattus no
C 39	17.8	84.8	243511	14 AC119076	AC119076 Rattus no
C 40	17.8	84.8	263637	14 AC094254	AC094254 Rattus no
C 41	17.8	84.8	319450	14 AC153281	AC153281 Bos tauru
C 42	17.4	82.9	421	6 AX408360	AX408360 Sequence
C 43	17.4	82.9	1766	6 BD248840	BD248840 47 human
C 44	17.4	82.9	2350	8 AY328894	AY328894 Homo sapi
C 45	17.4	82.9	2612	9 AF501309	AF501309 Mus muscu

ALIGNMENTS

RESULT 1
AF30557285/c
LOCUS AF30557285 579 bp DNA linear PRI 26-DEC-2000
DEFINITION Homo sapiens mucopolipin 1 (MCOLN1) gene, exons 6 and 7.
ACCESSION AF305576
VERSION AF305576.1 GI:11991202
KEYWORDS
SEGMENT
SOURCE 5 of 8
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 579)
AUTHORS Bargal, R., Avidan, N., Ben-Asher, E., Olender, Z., Zeigler, M.,
Frumkin, A., Raas-Rothschild, A., Glusman, G., Lancet, D. and Bach, G.
TITLE Identification of the gene causing mucopolipidosis type IV
JOURNAL Nat. Genet. 26 (1), 118-123 (2000)
PUBMED 10973263
REFERENCE 2 (bases 1 to 579)
AUTHORS Bargal, R., Avidan, N., Ben-Asher, E., Olender, A., Zeigler, M.,
Frumkin, A., Raas-Rothschild, A., Glusman, G., Lancet, D. and Bach, G.
TITLE Direct Submission
JOURNAL Submitted (13-SEP-2000) Molecular Genetics, The Weizmann Institute
of Science, P. O. Box 26, Rehovot 76100, Israel
FEATURES
source
1..579
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
129..225
/gene="MCOLN1"
/number=6
423..522
/gene="MCOLN1"
/number=7
exon
exon
ORIGIN
Query Match 100.0%; Score 21; DB 8; Length 579;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 15:55:50 ; Search time 2348.47 Seconds
(without alignments)
418.369 Million cells updates/sec

Title: US-10-754-446-4

Perfect score: 21
Sequence: 1 ctacacgtgctggaagacac 21

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_hc.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_est7.*
9: gb_gss1.*
10: gb_gss2.*
11: gb_gss3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	21	100.0	868	3	BI909885 603070517
C 2	18.4	87.6	225	8	CX226198 MBM01462
C 3	18.4	87.6	331	8	Z43611 HSC1GD031 n
C 4	18.4	87.6	333	8	T35364 EST83466 Hu
C 5	18.4	87.6	350	8	Z43604 HSC1GA071 n
C 6	18.4	87.6	423	8	H64102 YF57E02.r1
C 7	18.4	87.6	582	3	BP196543 BP196543
C 8	18.4	87.6	818	2	EG831705
C 9	18.4	87.6	826	2	EG913075
C 10	18.4	87.6	873	6	CD558014
C 11	18.4	87.6	878	5	BQ956452
C 12	18.4	87.6	901	5	BX433310
C 13	18.4	87.6	940	5	BQ858168
C 14	18.4	87.6	979	5	BQ918296
C 15	18.4	87.6	993	2	BG739714
C 16	18.4	87.6	1078	5	BX362786
C 17	18.4	87.6	1273	4	CR616477
C 18	18.4	87.6	1743	10	AV410283
C 19	18.4	87.6	1996	4	CR622331
C 20	18.4	87.6	3287	4	AK048278
C 21	18	85.7	901	7	CK159092
C 22	17.8	84.8	539	7	CR927955

C 23	17.8	84.8	552	3	BI775909
C 24	17.8	84.8	604	7	CN791999
C 25	17.8	84.8	660	9	BZ268910
C 26	17.8	84.8	1245	6	CA491046
C 27	17.8	84.8	1658	2	BE962361
C 28	17.4	82.9	219	3	BM293485
C 29	17.4	82.9	223	1	AI610095
C 30	17.4	82.9	271	8	N75483
C 31	17.4	82.9	276	1	AA330598
C 32	17.4	82.9	283	6	CA880163
C 33	17.4	82.9	308	1	AA879907
C 34	17.4	82.9	343	2	BB793665
C 35	17.4	82.9	364	8	H26993
C 36	17.4	82.9	375	5	BU735335
C 37	17.4	82.9	384	8	H42449
C 38	17.4	82.9	388	1	AI689890
C 39	17.4	82.9	395	2	BE690638
C 40	17.4	82.9	403	11	DE136711
C 41	17.4	82.9	409	1	AI190915
C 42	17.4	82.9	415	1	AI061413
C 43	17.4	82.9	418	1	AW172705
C 44	17.4	82.9	421	1	AA430026
C 45	17.4	82.9	421	8	N62729

ALIGNMENTS

RESULT 1
LOCUS BI909885/c
DEFINITION 603070517F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5219237 5', linear EST 16-OCT-2001 mRNA sequence.
ACCESSION BI909885
VERSION BI909885.1 GI:16173182
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 868)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LHAM11551 row: b column: 06
High quality sequence stop: 819.
Location/Qualifiers
1. .868
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5219237"
/tissue_type="leukocyte"
/lab_host="DH10B"
/clone_lib="NIH_MGC_118"
/note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note:

BI909885 868 bp mRNA linear EST 16-OCT-2001
603070517F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5219237 5',
mRNA sequence.
BI909885
BI909885.1 GI:16173182
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 868)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LHAM11551 row: b column: 06
High quality sequence stop: 819.
Location/Qualifiers
1. .868
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5219237"
/tissue_type="leukocyte"
/lab_host="DH10B"
/clone_lib="NIH_MGC_118"
/note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note:

FEATURES
source

```

this is a NIH_MGC Library."

ORIGIN
Query Match      100.0%; Score 21; DB 3; Length 868;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTCACCGTGTGGAGACACT 21
    |||||
Db 122 CTCACCGTGTGGAGACACT 102

RESULT 2
CX226198      225 bp  mRNA  linear  EST 29-DEC-2004
LOCUS      MBM01462 Mus Musculus hematopoietic BM-HPC5 cDNA library Mus
DEFINITION      musculus cDNA 5', mRNA sequence.
ACCESSION      CX226198
VERSION      CX226198.1 GI:56981490
KEYWORDS      EST.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 225)
AUTHORS      Williams,C., Wirta,V., Richter,K., Karlsson,C., Lundberg,J. and
              Carlsson,L.
TITLE      Expressed sequence tags of cDNA clones from a hematopoietic stem
              cell line expressing Lhx2
JOURNAL      Unpublished (2005)
COMMENT      Contact: Williams, C.
              Molecular Biotechnology
              Institution of Biotechnology
              Albatova University Center, KTH-Royal Institute of Technology, 106
              91 Stockholm, Sweden
              Tel: +46855378332.
              Fax: +46855378481
              Email: cecilia.williams@biotech.kth.se
              Seq primer: M13REV.
FEATURES
    source
    1..225
        /organism="Mus musculus"
        /mol_type="mRNA"
        /strain="C57BL/6-cast"
        /db_xref="taxon:10090"
        /sex="Male and Female"
        /tissue_type="Bone Marrow"
        /cell_type="Hematopoietic progenitor/stem cells
        immortalized by Lhx2"
        /cell_line="hematopoietic stem cell-like cell line
        BM-HPC5"
        /dev_stage="adult"
        /clone_lib="Mus Musculus hematopoietic BM-HPC5 cDNA
        library"
        /notes="Organ: Bone Marrow; Vector: pCMVSPORT6.1;
        Preamplified custom cDNA library by Invitrogen/ResGen"

ORIGIN
Query Match      87.6%; Score 18.4; DB 8; Length 225;
Best Local Similarity 95.0%; Pred. No. 4.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTCACCGTGTGGAGACAC 20
    |||||
Db 141 CTCACCATGTGGAGACAC 122

RESULT 3
Z43611
LOCUS      HSC1GD031 normalized infant brain cDNA Homo sapiens cDNA clone
DEFINITION      c-1gd03, mRNA sequence.

this is a NIH_MGC Library."

ORIGIN
Query Match      100.0%; Score 21; DB 3; Length 868;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTCACCGTGTGGAGACACT 21
    |||||
Db 122 CTCACCGTGTGGAGACACT 102

RESULT 2
CX226198      225 bp  mRNA  linear  EST 29-DEC-2004
LOCUS      MBM01462 Mus Musculus hematopoietic BM-HPC5 cDNA library Mus
DEFINITION      musculus cDNA 5', mRNA sequence.
ACCESSION      CX226198
VERSION      CX226198.1 GI:56981490
KEYWORDS      EST.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 225)
AUTHORS      Williams,C., Wirta,V., Richter,K., Karlsson,C., Lundberg,J. and
              Carlsson,L.
TITLE      Expressed sequence tags of cDNA clones from a hematopoietic stem
              cell line expressing Lhx2
JOURNAL      Unpublished (2005)
COMMENT      Contact: Williams, C.
              Molecular Biotechnology
              Institution of Biotechnology
              Albatova University Center, KTH-Royal Institute of Technology, 106
              91 Stockholm, Sweden
              Tel: +46855378332.
              Fax: +46855378481
              Email: cecilia.williams@biotech.kth.se
              Seq primer: M13REV.
FEATURES
    source
    1..225
        /organism="Mus musculus"
        /mol_type="mRNA"
        /strain="C57BL/6-cast"
        /db_xref="taxon:10090"
        /sex="Male and Female"
        /tissue_type="Bone Marrow"
        /cell_type="Hematopoietic progenitor/stem cells
        immortalized by Lhx2"
        /cell_line="hematopoietic stem cell-like cell line
        BM-HPC5"
        /dev_stage="adult"
        /clone_lib="Mus Musculus hematopoietic BM-HPC5 cDNA
        library"
        /notes="Organ: Bone Marrow; Vector: pCMVSPORT6.1;
        Preamplified custom cDNA library by Invitrogen/ResGen"

ORIGIN
Query Match      87.6%; Score 18.4; DB 8; Length 225;
Best Local Similarity 95.0%; Pred. No. 4.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTCACCGTGTGGAGACAC 20
    |||||
Db 141 CTCACCATGTGGAGACAC 122

RESULT 3
Z43611
LOCUS      HSC1GD031 normalized infant brain cDNA Homo sapiens cDNA clone
DEFINITION      c-1gd03, mRNA sequence.

```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 14:15:14 ; Search time 284.928 Seconds
(without alignments)
491.207 Million cell updates/sec

Title: US-10-754-446-4

Perfect score: 21

Sequence: 1 ctcaccgtgtggaagacact 21

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21.*

- 1: geneseqn1980s.*
- 2: geneseqn1990s.*
- 3: geneseqn2000s.*
- 4: geneseqn2001as.*
- 5: geneseqn2001bs.*
- 6: geneseqn2002as.*
- 7: geneseqn2002bs.*
- 8: geneseqn2003as.*
- 9: geneseqn2003bs.*
- 10: geneseqn2003cs.*
- 11: geneseqn2003ds.*
- 12: geneseqn2004as.*
- 13: geneseqn2004bs.*
- 14: geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	21	14 AEB28960	Aeb28960 Human MCO
C 2	21	100.0	11365	4 AAK73827	Aak73827 Human imm
C 3	21	100.0	13270	9 ADB84283	Adb84283 Human muc
C 4	21	100.0	20046	4 AAK73826	Aak73826 Human imm
5	20	95.2	20	9 ADB84282	Adb84282 Human muc
6	18.4	87.6	1619	4 AAI61022	Aai61022 Human pol
C 7	18.4	87.6	1740	6 AAI17100	Aai17100 Human TRP
C 8	18.4	87.6	1740	6 ABL40755	Abi40755 Human TLC
C 9	18.4	87.6	1741	4 AAI59236	Aai59236 Human pol
C 10	18.4	87.6	1743	9 ADB84284	Adb84284 Human muc
C 11	18.4	87.6	1743	14 ADV66231	Adv66231 TRP-like
C 12	18.4	87.6	2051	13 ADP25150	Adp25150 PRO polyp
C 13	18.4	87.6	2052	4 AAF81753	Aaf81753 Human mem
C 14	18.4	87.6	2092	6 ABL90358	Abi90358 Human pol
C 15	18.4	87.6	2094	3 AAA39067	Aaa39067 Human sec
C 16	18.4	87.6	2095	6 AAI71699	Aai71699 Human TRP
C 17	18.4	87.6	2095	6 ABL40754	Abi40754 Human TLC
C 18	18.4	87.6	2095	14 ADV66229	Adv66229 TRP-like
C 19	18.4	87.6	2140	12 ADQ24925	Adq24925 Human sof

C	20	17.4	82.9	370	5	ABA12621	Abal2621 Human ner
C	21	17.4	82.9	421	6	ABN94509	Abn94509 Gene #100
C	22	17.4	82.9	1766	3	AAA78381	Aaa78381 Human sec
C	23	17.4	82.9	1766	8	ADA39925	Ada39925 Human sec
C	24	17.4	82.9	1766	8	ACC50509	Acc50509 Human sec
C	25	17.4	82.9	1766	8	ABZ71282	Abz71282 Secreted
C	26	17.4	82.9	1766	9	ADB91188	Adb91188 Human sec
C	27	17.4	82.9	1766	10	ADC73569	Adc73569 Human sec
C	28	17.4	82.9	1766	10	ADA56112	Ada56112 Gene enco
C	29	17.4	82.9	2530	14	AEA47619	Aea47619 Nucleotid
C	30	17.4	82.9	2750	6	ABL92130	Abi92130 Human Tum
C	31	17.4	82.9	2750	10	ABX72055	Abx72055 DNA enco
C	32	17.4	82.9	3258	4	AAH42269	Aah42269 Nucleotid
C	33	17.4	82.9	3263	10	ADA53072	Ada53072 Human cod
C	34	17.4	82.9	3272	10	ADE31397	Ade31397 Human dia
C	35	17.4	82.9	3351	11	ACN44967	Acn44967 Human mrn
C	36	17.4	82.9	125534	11	ACN44966	Acn44966 Human gen
C	37	17	81.0	2092	5	AA572274	Aae72274 DNA enco
C	38	16.8	80.0	102	10	ADG25870	Adg25870 INPIONCHO
C	39	16.8	80.0	297	8	ABX43933	Abx43933 Bovine ES
C	40	16.8	80.0	399	8	ABX45303	Abx45303 Bovine ES
C	41	16.8	80.0	433	8	ABX39870	Abx39870 Bovine ES
C	42	16.8	80.0	472	6	ABN73626	Abn73626 Bovine em
C	43	16.8	80.0	586	10	ADD34724	Add34724 Mouse mit
C	44	16.8	80.0	594	12	ACH77455	Ach77455 Human gen
C	45	16.8	80.0	2988	8	ACA57290	Aca57290 Human adi

ALIGNMENTS

RESULT 1

AEB28960
ID AEB28960 standard; DNA; 21 BP.

AC AEB28960;

XX 22-SEP-2005 (first entry)

XX Human MCOLN1 gene PCR primer PRI R2 SEQ ID NO:4.

XX mucolipidosis IV; high throughput screening; mucolipin-1; MCOLN1; PCR; primer; ss.

XX Homo sapiens.

XX US2005153300-A1.

XX 14-JUL-2005.

XX 09-JAN-2004; 2004US-00754446.

XX 09-JAN-2004; 2004US-00754446.

XX (QUES-) QUEST DIAGNOSTICS INC.

XX Sun W, Hantash F;

XX WPI; 2005-521160/53.

XX Diagnosing mucolipidosis (ML) IV mutant sequence in nucleic acids, by amplifying the nucleic acid, detecting amplified product with labeled oligonucleotide probes via a change in fluorescence which indicates the presence of an ML IV mutant.

XX Claim 1; SEQ ID NO 4; 15pp; English.

XX The invention relates to a method (M1) for detecting the presence of mucolipidosis IV (ML IV) mutant sequences in a nucleic acid. (M1) involves contacting the nucleic acid with oligonucleotide primers and probes, conducting amplification by temperature cycling and monitoring the accumulation of amplified nucleic acid by detecting an increase in donor fluorophore fluorescence or decrease in acceptor fluorophore

CC fluorescence, which indicates the presence of ML IV mutant sequence. Also
 CC described: (1) an oligonucleotide (ON1) comprising a sequence
 CC complementary to the coding or non-coding strand of the mucolin-1
 CC (MCOLN1) gene; and (2) a kit (K1) for amplifying sequences of ML IV
 CC sequences, comprising ON1. (M1) is useful for determining the presence of
 CC ML IV mutant sequence in a nucleic acid, and/or wild-type sequences. (M1)
 CC is useful for genetic counseling of individuals at risk for ML IV or at
 CC risk to having offspring afflicted with ML IV. The MCOLN1 gene is located
 CC on chromosome 19p13. The present sequence represents a PCR primer for the
 CC MCOLN1 gene, which is used in the exemplification of the present
 CC invention.

XX SQ Sequence 21 BP; 5 A; 7 C; 5 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 21; DB 14; Length 21;

Best Local Similarity 100.0%; Pred. No. 2;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTCACCGTGTGGAGACACT 21

Db 1 CTCACCGTGTGGAGACACT 21

RESULT 2

AAK73827/c

ID AAK73827 standard; DNA; 11365 BP.

XX AC AAK73827;

XX DT 07-NOV-2001 (first entry)

XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:28639.

XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

XX KW Cytostatic; gene therapy; vaccine; metastasis; ds.

XX OS Homo sapiens.

XX PN WO200157182-A2.

XX PD 09-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001354.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 24-FEB-2000; 2000US-0184664P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-0190076P.

XX PR 18-APR-2000; 2000US-0198123P.

XX PR 19-MAY-2000; 2000US-0205515P.

XX PR 07-JUN-2000; 2000US-0209467P.

XX PR 28-JUN-2000; 2000US-0214886P.

XX PR 30-JUN-2000; 2000US-0215135P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 07-JUL-2000; 2000US-0216880P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 11-JUL-2000; 2000US-0217496P.

XX PR 14-JUL-2000; 2000US-0218290P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 26-JUL-2000; 2000US-0220964P.

XX PR 14-AUG-2000; 2000US-0224518P.

XX PR 14-AUG-2000; 2000US-0224519P.

XX PR 14-AUG-2000; 2000US-0225213P.

XX PR 14-AUG-2000; 2000US-0225214P.

XX PR 14-AUG-2000; 2000US-0225266P.

XX PR 14-AUG-2000; 2000US-0225267P.

XX PR 14-AUG-2000; 2000US-0225268P.

XX PR 14-AUG-2000; 2000US-0225270P.

XX PR 14-AUG-2000; 2000US-0225447P.

XX PR 14-AUG-2000; 2000US-0225757P.

XX PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 13-OCT-2000; 2000US-0239335P.
 PR 13-OCT-2000; 2000US-0239337P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 16:41:33 ; Search time 507.192 Seconds
(without alignments)
342.389 Million cell updates/sec

Title: US-10-754-446-4

Perfect score: 21

Sequence: 1 ctccaccgtgctggaagacact 21

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main:

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	21	9	US-10-754-446-4
2	21	100.0	13270	3	US-09-851-494B-1
3	20	95.2	20	3	US-09-851-494B-11
4	18.4	87.6	25	10	US-11-060-756-194946
5	18.4	87.6	1400	10	US-11-060-756-2850
6	18.4	87.6	1400	10	US-11-060-756-7122
7	18.4	87.6	1740	3	US-09-828-466-3
8	18.4	87.6	1740	5	US-10-103-458-3
9	18.4	87.6	1743	8	US-10-782-695-6
10	18.4	87.6	2051	3	US-09-851-494B-2
11	18.4	87.6	2052	3	US-09-985-529-50
12	18.4	87.6	2052	3	US-09-969-680A-50
13	18.4	87.6	2052	10	US-11-048-692-50
14	18.4	87.6	2092	6	US-10-264-237-920
15	18.4	87.6	2094	3	US-09-820-893-26
16	18.4	87.6	2094	7	US-10-607-565-26
17	18.4	87.6	2095	3	US-09-828-466-1
18	18.4	87.6	2095	5	US-10-103-458-1
19	18.4	87.6	2095	8	US-10-782-695-4
20	18.4	87.6	2140	8	US-10-723-860-7745
21	17.4	82.9	421	3	US-09-880-107-1007
22	17.4	82.9	644	5	US-10-027-632-193464
23	17.4	82.9	644	6	US-10-027-632-193464

C 24	17.4	82.9	1766	3	US-09-895-238-11	Sequence 11, Appl
C 25	17.4	82.9	1766	8	US-10-885-039-11	Sequence 11, Appl
C 26	17.4	82.9	1766	9	US-10-472-533-103	Sequence 103, App
C 27	17.4	82.9	2530	3	US-10-986-466-2	Sequence 2, Appl
C 28	17.4	82.9	2750	3	US-09-918-715-282	Sequence 282, App
C 29	17.4	82.9	2750	8	US-10-474-794-282	Sequence 282, App
C 30	17.4	82.9	2750	9	US-10-979-159-282	Sequence 282, App
C 31	17.4	82.9	3263	6	US-10-094-749-640	Sequence 640, App
C 32	17.4	82.9	3351	5	US-10-087-192-1679	Sequence 1679, Ap
C 33	17.4	82.9	125534	5	US-10-087-192-1678	Sequence 1678, Ap
C 34	17	81.0	2092	3	US-10-450-763-8078	Sequence 8078, Ap
C 35	16.8	80.0	297	3	US-09-960-352-9098	Sequence 9098, Ap
C 36	16.8	80.0	392	7	US-10-424-589-58115	Sequence 58115, A
C 37	16.8	80.0	399	3	US-09-960-352-10468	Sequence 10468, A
C 38	16.8	80.0	433	3	US-09-960-352-5035	Sequence 5035, Ap
C 39	16.8	80.0	472	3	US-09-876-143-565	Sequence 565, App
C 40	16.8	80.0	594	6	US-10-029-386-10650	Sequence 10650, A
C 41	16.8	80.0	600	9	US-10-972-079-81475	Sequence 81475, A
C 42	16.8	80.0	600	9	US-10-972-079-81476	Sequence 81476, A
C 43	16.8	80.0	600	9	US-10-972-079-81477	Sequence 81477, A
C 44	16.8	80.0	890	7	US-10-437-963-94453	Sequence 94453, A
C 45	16.8	80.0	462586	9	US-10-476-264-420	Sequence 420, App

ALIGNMENTS

RESULT 1
US-10-754-446-4
; Sequence 4, Application US/10754446
; Publication No. US20050153300A1
; GENERAL INFORMATION:
; APPLICANT: SUN, WEIMIN
; APPLICANT: HANTASH, FERAS
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DETECTION OF
; FILE REFERENCE: 034827-2301
; CURRENT APPLICATION NUMBER: US/10/754, 446
; CURRENT FILING DATE: 2004-01-09
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 3.2
; SEQ ID NO 4
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-10-754-446-4

Query Match 100.0%; Score 21; DB 9; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTACCCGTGCTGGAAGACACT 21
| | | | | | | | | | | | | | | | | | | | | |
Db 1 CTACCCGTGCTGGAAGACACT 21

RESULT 2
US-09-851-494B-1/c
; Sequence 1, Application US/09851494B
; Publication No. US20030064363A1
; GENERAL INFORMATION:
; APPLICANT: ML4 Foundation
; APPLICANT: Goldin, Ehud
; APPLICANT: Staugenhaupt, Susan A.
; APPLICANT: Sun, Mei
; APPLICANT: Aclerno, James S.
; TITLE OF INVENTION: A Gene Encoding A New TRP Channel is Mutated in Mucopolidosis IV
; FILE REFERENCE: 3394/1H557US1
; CURRENT APPLICATION NUMBER: US/09/851, 494B
; CURRENT FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 11

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 16:59:57 ; Search time 258.384 Seconds
(without alignments)
65.779 Million cell updates/sec

Title: US-10-754-446-4

Perfect score: 21

Sequence: 1 ctcaccgtgtggaagacacat 21

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 6038814 seqs, 404674181 residues

Total number of hits satisfying chosen parameters: 12077628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA New:*
1: /cgm2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
2: /cgm2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /cgm2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /cgm2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
5: /cgm2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
6: /cgm2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
7: /cgm2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
8: /cgm2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2.*
9: /cgm2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3.*
10: /cgm2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16.2	77.1	77246	7	US-11-124-368A-2907
2	15.8	75.2	32157	6	US-10-995-561-13352
3	15.4	73.3	19	8	US-11-101-244-151656
4	15.4	73.3	1238	6	US-11-083-784-151656
5	15.4	73.3	1238	6	US-10-821-234-1
6	15.4	73.3	1244	7	US-11-000-688-733
7	15.4	73.3	6990	7	US-11-000-688-609
8	15.4	73.3	175673	7	US-11-121-086-55
9	15.2	72.4	21	6	US-10-310-914A-1157477
10	15.2	72.4	579	7	US-11-128-061-342
11	15.2	72.4	579	7	US-11-128-061-3984
12	15.2	72.4	866	6	US-10-750-185-61786
13	15.2	72.4	866	6	US-10-750-185-61786
14	15.2	72.4	1008	6	US-10-750-185-54439
15	15.2	72.4	1008	6	US-10-750-623-54439
16	15.2	72.4	1077	6	US-10-517-939-369
17	15.2	72.4	1307	6	US-10-750-185-27270
18	15.2	72.4	1307	6	US-10-750-623-27270
19	15.2	72.4	1350	7	US-11-055-822-289
20	15.2	72.4	3272	6	US-10-750-185-24646
21	15.2	72.4	3272	6	US-10-750-623-24646
22	15.2	72.4	4730	7	US-11-128-061-520
23	15.2	72.4	5595	6	US-10-955-054A-11

C 24	15.2	72.4	154452	7	US-11-121-086-74	Sequence 74, Appl
C 25	15	71.4	201	6	US-10-995-561-26690	Sequence 26690, A
C 26	15	71.4	201	6	US-10-995-561-73444	Sequence 73444, A
C 27	15	71.4	14082	6	US-10-995-561-13445	Sequence 13445, A
C 28	15	71.4	305312	6	US-10-995-561-13236	Sequence 13236, A
C 29	14.8	70.5	19	8	US-11-101-244-285326	Sequence 285326, A
C 30	14.8	70.5	19	9	US-11-083-784-285326	Sequence 285326, A
C 31	14.8	70.5	25	7	US-11-121-849-182614	Sequence 182614, A
C 32	14.8	70.5	201	6	US-10-995-561-23969	Sequence 23969, A
C 33	14.8	70.5	327	7	US-11-000-688-786	Sequence 786, App
C 34	14.8	70.5	1105	6	US-11-000-463-581	Sequence 581, App
C 35	14.8	70.5	1296	6	US-10-525-710-23	Sequence 23, Appl
C 36	14.8	70.5	1388	7	US-11-128-061-969	Sequence 969, App
C 37	14.8	70.5	1481	7	US-11-000-688-788	Sequence 788, App
C 38	14.8	70.5	1646	6	US-10-750-185-55975	Sequence 55975, A
C 39	14.8	70.5	1646	6	US-10-750-623-55975	Sequence 55975, A
C 40	14.8	70.5	2244	7	US-11-128-061-528	Sequence 528, App
C 41	14.8	70.5	98345	7	US-11-112-908-36	Sequence 36, Appl
C 42	14.8	70.5	127340	7	US-11-112-908-35	Sequence 35, Appl
C 43	14.8	70.5	207835	7	US-11-121-086-39	Sequence 39, Appl
C 44	14.8	70.5	207835	7	US-11-121-086-40	Sequence 40, Appl
C 45	14.6	69.5	25	7	US-11-136-527-319963	Sequence 319963, A

ALIGNMENTS

RESULT 1

US-11-124-368A-2907
; Sequence 2907, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; PRIOR FILING DATE: 2005-05-09
; PRIOR FILING DATE: 2004-05-07
; PRIOR FILING DATE: 2004-05-07
; PRIOR FILING DATE: 2004-11-09
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2907
; LENGTH: 77246
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 10718, 10719, 10720, 10721, 10722, 10723, 10724, 10725, 10726,
; LOCATION: 10727, 10728, 10729, 10730, 10731, 10732, 10733, 10734,
; LOCATION: 10735, 10736, 10737, 10738, 10739, 10740, 10741, 10742,
; LOCATION: 10743, 10744, 10745, 10746, 10747, 10748, 10749, 10750
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 10751, 10752, 10753, 10754, 10755, 10756, 10757, 10758, 10759,
; LOCATION: 10760, 10761, 10762, 10763, 10764, 10765, 10766, 10767,
; LOCATION: 10768, 10769, 10770, 10771, 10772, 10773, 10774, 10775,
; LOCATION: 10776, 10777, 10778, 10779, 10780, 10781, 10782, 10783
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 10784, 10785, 10786, 10787, 10788, 10789, 10790, 10791, 10792,
; LOCATION: 10793, 10794, 10795, 10796, 10797, 10798, 10799, 10800,
; LOCATION: 10801, 10802, 10803, 10804, 10805, 10806, 10807, 10808,
; LOCATION: 10809, 10810, 10811, 10812, 10813, 10814, 10815, 10816
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature

```

; LOCATION: 10817, 10818, 10819, 10820, 10821, 10822, 10823, 10824, 10825,
; LOCATION: 10826, 10827, 10828, 10829, 10830, 10831, 10832, 10833,
; LOCATION: 10834, 10835, 10836, 10837, 10838, 10839, 10840, 10841,
; LOCATION: 10842, 10843, 10844, 10845, 10846, 10847, 10848, 10849
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 10850, 10851, 10852, 10853, 10854, 10855, 10856, 10857, 10858,
; LOCATION: 10859, 10860, 10861, 10862, 10863, 10864, 10865, 10866,
; LOCATION: 10867, 10868, 10869, 10870, 10871, 10872, 10873, 10874,
; LOCATION: 10875, 10876, 10877, 10878, 10879, 10880, 10881, 10882
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 10883, 10884, 10885, 10886, 10887, 10888, 10889, 10890, 10891,
; LOCATION: 10892, 10893, 10894, 10895, 10896, 10897, 10898, 10899,
; LOCATION: 10900, 10901, 10902, 10903, 10904, 10905, 10906, 10907,
; LOCATION: 10908, 10909, 10910, 10911, 10912, 10913, 10914, 10915
; OTHER INFORMATION: n = A,T,C or G
US-11-124-368A-2907

Query Match          77.1%; Score 16.2; DB 7; Length 77246;
Best Local Similarity 85.7%; Pred. No. 77;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTCACCGTGTGGGAAGACACT 21
Db 633 CTCACCGTGCAGGAGACACT 653

RESULT 2
US-10-995-561-13352/c
; Sequence 13352, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995.561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13352
; LENGTH: 32157
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13352

Query Match          75.2%; Score 15.8; DB 6; Length 32157;
Best Local Similarity 89.5%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TCACCGTGTGGGAAGACAC 20
Db 17401 TCACCGTGTGGGAAGAAC 17383

RESULT 3
US-11-101-244-151656
; Sequence 151656, Application US/11010244
; Publication No. US20050246794A1
; GENERAL INFORMATION:
; APPLICANT: Dharmacon, Inc.
; APPLICANT: Khvorova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scaringe, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCE: 134990S
; CURRENT APPLICATION NUMBER: US/11/01.244

```

```

; CURRENT FILING DATE: 2005-04-07
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary
; SEQ ID NO 151656
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-101-244-151656

```

```

Query Match          73.3%; Score 15.4; DB 8; Length 19;
Best Local Similarity 76.5%; Pred. No. 1.3e+02;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 2 TCACCGTGTGGGAAGAC 18
Db 1 UCAACGUGUGGAAGAC 17

```

```

RESULT 4
US-11-083-784-151656
; Sequence 151656, Application US/11083784
; Publication No. US20050245475A1
; GENERAL INFORMATION:
; APPLICANT: Dharmacon, Inc.
; APPLICANT: Khvorova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scaringe, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCE: 134990S
; CURRENT APPLICATION NUMBER: US/11/083,784
; CURRENT FILING DATE: 2005-03-18
; PRIOR APPLICATION NUMBER: US/10/714,333
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary
; SEQ ID NO 151656
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-083-784-151656

```

```

Query Match          73.3%; Score 15.4; DB 9; Length 19;
Best Local Similarity 76.5%; Pred. No. 1.3e+02;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 2 TCACCGTGTGGGAAGAC 18
Db 1 UCAACGUGUGGAAGAC 17

```

```

RESULT 5
US-10-821-234-1
; Sequence 1, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Grain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07

```


GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 16:32:47 ; Search time 76.44 Seconds
(without alignments)
488.341 Million cell updates/sec

Title: US-10-754-446-4

Perfect score:

Sequence: 1 ctcacgctgctggaagacact 21

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 200000000
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

```
Database : Issued Patents_NA.*
1 /cgn2_6/ptodata/1/ina/1_COMB.seq.*
2 /cgn2_6/ptodata/1/ina/5_COMB.seq.*
3 /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4 /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5 /cgn2_6/ptodata/1/ina/H_COMB.seq.*
6 /cgn2_6/ptodata/1/ina/PCRUS_COMB.seq.*
7 /cgn2_6/ptodata/1/ina/PP_COMB.seq.*
8 /cgn2_6/ptodata/1/ina/RE_COMB.seq.*
9 /cgn2_6/ptodata/1/ina/backfiles1.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	21	100.0	601	3	US-09-949-016-59085	Sequence 59085, A
C 2	21	100.0	15353	3	US-09-949-016-13454	Sequence 13454, A
C 3	18.4	87.6	2051	3	US-09-949-016-1712	Sequence 1712, Ap
C 4	16.2	77.1	25	3	US-09-396-196G-50417	Sequence 50417, A
C 5	16.2	77.1	471	3	US-09-270-767-1970	Sequence 1970, Ap
C 6	16.2	77.1	471	3	US-09-270-767-17252	Sequence 17252, A
C 7	16.2	77.1	601	3	US-09-949-016-178570	Sequence 178570, A
C 8	16.2	77.1	601	3	US-09-949-016-178571	Sequence 178571, A
C 9	16.2	77.1	797	3	US-09-270-767-7266	Sequence 7266, Ap
C 10	16.2	77.1	797	3	US-09-270-767-22548	Sequence 22548, A
C 11	16.2	77.1	3385	3	US-09-509-800-1	Sequence 1, Appli
C 12	16.2	77.1	12470	3	US-09-949-002-847	Sequence 847, App
C 13	16.2	77.1	17896	3	US-09-949-016-16867	Sequence 16867, A
C 14	16.2	77.1	32665	3	US-09-949-016-14098	Sequence 14098, A
C 15	16.2	77.1	52667	3	US-09-949-016-12019	Sequence 12019, A
C 16	16.2	77.1	141560	3	US-09-949-016-16476	Sequence 16476, A
C 17	15.8	75.2	25	3	US-09-396-196G-50406	Sequence 50406, A
C 18	15.8	75.2	659	3	US-09-533-559-6889	Sequence 6889, Ap
C 19	15.8	75.2	1293	3	US-09-902-540-5426	Sequence 5426, Ap
C 20	15.8	75.2	24154	3	US-09-949-016-16374	Sequence 16374, A
C 21	15.8	75.2	24655	3	US-09-949-016-15561	Sequence 15561, A
C 22	15.8	75.2	34953	3	US-09-902-540-1263	Sequence 1263, Ap
C 23	15.4	73.3	358	3	US-09-513-999C-3422	Sequence 3422, Ap
C 24	15.4	73.3	601	3	US-09-949-016-51396	Sequence 51396, A

ALIGNMENTS

RESULT 1

```

US-09-949-016-59085/c
/ Sequence 59085, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN
/ COMPOSITIONS WITH HUMAN DISEASE, N
/ FILE REFERENCE: CLO01307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: PASCSEQ for Windows Version 4.0
/ SEQ ID NO 59085
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-59085

```

```
Query Match      100.0%; Score 21; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 0.99;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 CTACCGTGCTGGAAGACACT 21
|||
Db 504 CTACCGTGCTGGAAGACACT 484

RESULT. T. 2

RES-001.2
US-09-949-016-13454/c
; Sequence 13454, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20

```
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13454
; LENGTH: 15353
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13454

Query Match      100.0%; Score 21; DB 3; Length 15353;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCACCGTGTGGAGACACT 21
Db 7637 CTCACCGTGTGGAGACACT 7617

RESULT 3
US-09-949-016-1712/c
; Sequence 1712, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT FILING DATE: 2000-04-14
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1712
; LENGTH: 2051
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1712

Query Match      87.6%; Score 18.4; DB 3; Length 2051;
Best Local Similarity 95.0%; Pred. No. 20;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TCACCGTGTGGAGACACT 21
Db 1006 TCACCGTGTGGAGACACT 987

RESULT 4
US-09-396-196G-50417
; Sequence 50417, Application US/09396196G
; Patent No. 6821724
; GENERAL INFORMATION:
; APPLICANT: Michael Mittmann
; APPLICANT: David Mack
; APPLICANT: David Lockhart
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis
; FILE REFERENCE: 3101.1
; CURRENT APPLICATION NUMBER: US/09/396,196G
; CURRENT FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,678
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 127806
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50417
```

```
; LENGTH: 25
; TYPE: DNA
; ORGANISM: mus musculus
US-09-396-196G-50417

Query Match      77.1%; Score 16.2; DB 3; Length 25;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTCACCGTGTGGAGACACT 21
Db 4 CACACCGTGTGGAGACACT 24

RESULT 5
US-09-270-767-1970/c
; Sequence 1970, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1970
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-1970

Query Match      77.1%; Score 16.2; DB 3; Length 471;
Best Local Similarity 85.7%; Pred. No. 2.1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTCACCGTGTGGAGACACT 21
Db 451 CTGACCGTGTGGAGACACT 431

RESULT 6
US-09-270-767-17252/c
; Sequence 17252, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17252
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-17252

Query Match      77.1%; Score 16.2; DB 3; Length 471;
Best Local Similarity 85.7%; Pred. No. 2.1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTCACCGTGTGGAGACACT 21
Db 451 CTGACCGTGTGGAGACACT 431

RESULT 7
US-09-949-016-178570
; Sequence 178570, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 15:49:05 ; Search time 982.24 Seconds
(without alignments)
1157.425 Million cell updates/sec

Title: US-10-754-446-3

Perfect score: 20

Sequence: 1 cttgctctgtgcccaggct 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sts.*
11: gb_sy.*
12: gb_un.*
13: gb_vl.*
14: gb_htg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	20	100.0	51	6 AX163311	Sequence
C 2	20	100.0	51	6 AX199337	Sequence
C 3	20	100.0	51	6 AX199338	Sequence
C 4	20	100.0	71	6 AX522607	Sequence
C 5	20	100.0	77	6 AX523183	Sequence
C 6	20	100.0	98	6 BD039678	Sequence
C 7	20	100.0	98	6 AX904145	Sequence
C 8	20	100.0	102	6 BD037012	Sequence
C 9	20	100.0	102	6 AX901479	Sequence
C 10	20	100.0	107	6 CQ685626	Sequence
C 11	20	100.0	108	6 CQ694139	Sequence
C 12	20	100.0	114	6 BD044993	Sequence
C 13	20	100.0	114	6 AX909460	Sequence
C 14	20	100.0	117	6 BD037394	Sequence
C 15	20	100.0	117	6 BD044139	Sequence
C 16	20	100.0	117	6 AX901861	Sequence
C 17	20	100.0	117	6 AX908606	Sequence
C 18	20	100.0	128	6 BD034714	Sequence

19	20	100.0	128	6 AX899181	Sequence
C 20	20	100.0	129	6 CQ708055	Sequence
21	20	100.0	145	6 BD035848	Sequence
22	20	100.0	145	6 BD045580	Sequence
C 23	20	100.0	145	6 BD118446	EST and e
C 24	20	100.0	145	6 AR422893	Sequence
25	20	100.0	145	6 AX900315	Sequence
26	20	100.0	145	6 AX910047	Sequence
C 27	20	100.0	145	6 AX983587	Sequence
C 28	20	100.0	147	6 CQ686625	Sequence
29	20	100.0	151	6 BD043797	Sequence
30	20	100.0	151	6 AX908264	Sequence
C 31	20	100.0	152	6 BD115706	EST and e
C 32	20	100.0	152	6 BD117639	Sequence
C 33	20	100.0	152	6 AR420153	Sequence
C 34	20	100.0	152	6 AR422086	Sequence
C 35	20	100.0	152	6 AX980847	Sequence
C 36	20	100.0	152	6 AX982780	Sequence
37	20	100.0	154	10 AB134303	Homo sapi
C 38	20	100.0	155	6 BD049251	Sequence
C 39	20	100.0	155	6 AX913718	Sequence
C 40	20	100.0	158	6 BD115744	EST and e
C 41	20	100.0	158	6 BD116463	EST and e
C 42	20	100.0	158	6 AR420191	Sequence
C 43	20	100.0	158	6 AR420910	Sequence
C 44	20	100.0	158	6 AX980885	Sequence
C 45	20	100.0	158	6 AX981604	Sequence

ALIGNMENTS

RESULT 1
AX163311/c
LOCUS AX163311 51 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 6639 from Patent WO0140521.
ACCESSION AX163311
VERSION AX163311.1 GI:14544642
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE 1
AUTHORS Shimkets,R.A. and Leach,M.
TITLE Nucleic acids containing single nucleotide polymorphisms and methods of use thereof
JOURNAL Patent: WO 0140521-A 6639 07-JUN-2001;
FEATURES
Location/Qualifiers
source 1..51
misc_feature 26
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="1 of 2 allelic variants (6640 is other entry)"
Accession number cg38821538"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 51;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTGCTCTGTGCCCCAGGCT 20
|||||
Db 37 CTTGCTCTGTGCCCCAGGCT 18

RESULT 2
AX199337/c
LOCUS AX199337 51 bp DNA linear PAT 29-AUG-2001
DEFINITION Sequence 267 from Patent WO0151670.

ACCESSION AX199337
VERSION AX199337.1 GI:15389718
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Shimkets,R.A. and Leach,M.D.
TITLE Nucleic acids containing single nucleotide polymorphisms and
methods of use thereof
JOURNAL Patent: WO 0151670-A 267 19-JUL-2001;
Curagen Corporation (US)
FEATURES Location/Qualifiers
source
1..51
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
misc_feature
26
/notes="1 of 2 allelic variants (268 is other entry)"
Accession number cg43011316"
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 51;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTTGCTCTGTGCCAGGCT 20
|||||
Db 49 CTTGCTCTGTGCCAGGCT 30
RESULT 3
AX199338/c
LOCUS AX199338 51 bp DNA linear PAT 29-AUG-2001
DEFINITION Sequence 268 from Patent WO0151670.
ACCESSION AX199338
VERSION AX199338.1 GI:15389719
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Shimkets,R.A. and Leach,M.D.
TITLE Nucleic acids containing single nucleotide polymorphisms and
methods of use thereof
JOURNAL Patent: WO 0151670-A 268 19-JUL-2001;
Curagen Corporation (US)
FEATURES Location/Qualifiers
source
1..51
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
misc_feature
26
/notes="2 of 2 allelic variants (267 is other entry)"
Accession number cg43011316"
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 51;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTTGCTCTGTGCCAGGCT 20
|||||
Db 49 CTTGCTCTGTGCCAGGCT 30
RESULT 4
AX522607
LOCUS AX522607 71 bp DNA linear PAT 24-OCT-2002

DEFINITION Sequence 277 from Patent WO02064731.
ACCESSION AX522607
VERSION AX522607.1 GI:24411561
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Telerman,A., Anson,R., Tuijnder,M. and Susini,L.
TITLE Sequences involved in phenomena of tumour suppression, tumour
reversion, apoptosis and/or virus resistance and their use as
medicines
JOURNAL Patent: WO 02064731-A 277 22-AUG-2002;
Molecular Engines Laboratories (FR)
FEATURES Location/Qualifiers
source
1..71
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 71;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTTGCTCTGTGCCAGGCT 20
|||||
Db 11 CTTGCTCTGTGCCAGGCT 30
RESULT 5
AX523183/c
LOCUS AX523183 77 bp DNA linear PAT 24-OCT-2002
DEFINITION Sequence 853 from Patent WO02064731.
ACCESSION AX523183
VERSION AX523183.1 GI:24412137
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Telerman,A., Anson,R., Tuijnder,M. and Susini,L.
TITLE Sequences involved in phenomena of tumour suppression, tumour
reversion, apoptosis and/or virus resistance and their use as
medicines
JOURNAL Patent: WO 02064731-A 853 22-AUG-2002;
Molecular Engines Laboratories (FR)
FEATURES Location/Qualifiers
source
1..77
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 77;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTTGCTCTGTGCCAGGCT 20
|||||
Db 39 CTTGCTCTGTGCCAGGCT 20
RESULT 6
BD039678/c
LOCUS BD039678 98 bp DNA linear PAT 27-AUG-2002
DEFINITION Sequence tag and encoded human protein.
ACCESSION BD039678
VERSION BD039678.1 GI:22581420

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 15:55:50 ; Search time 2236.64 Seconds
(without alignments)
418.369 Million cell updates/sec

Title: US-10-754-446-3

Perfect score: 20

Sequence: 1 cttgctgtgtgccaggct 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

1: gb_est1.*

2: gb_est2.*

3: gb_est3.*

4: gb_hic.*

5: gb_est4.*

6: gb_est5.*

7: gb_est6.*

8: gb_est7.*

9: gb_ges1.*

10: gb_ges2.*

11: gb_ges3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	20	100.0	38	1	AA054107 zf50f06.r
c 2	20	100.0	39	1	AA911358 oe76d10.s
c 3	20	100.0	46	9	AZ576121 AST-T3250
c 4	20	100.0	49	5	C20875 HUMS000494
c 5	20	100.0	50	1	AI708455 as97e06.x
c 6	20	100.0	50	1	AU102403 AU102403
c 7	20	100.0	52	1	AA903439 ok48h04.s
c 8	20	100.0	61	1	AI735552 at17e12.x
c 9	20	100.0	64	1	AW304187 xv61e11.x
c 10	20	100.0	68	9	AZ537585 AST-2P881
c 11	20	100.0	69	1	AA215549 zr97e12.r
c 12	20	100.0	70	1	AA604042 no44h11.s
c 13	20	100.0	82	2	BG236597 nai45g03.
c 14	20	100.0	84	9	AZ757146 ew06c09.r
c 15	20	100.0	101	1	AA262246 z826e09.s
c 16	20	100.0	102	6	CD640446
c 17	20	100.0	106	2	BF875905 RC3-ET013
c 18	20	100.0	108	2	BE141837 IL5-HT011
c 19	20	100.0	108	8	T58114 yb26d03.r1
c 20	20	100.0	111	2	BE246735 TCAP1D44
c 21	20	100.0	111	9	B66475 CIT-HSP-201
c 22	20	100.0	113	1	AW780390 ho05a12.x

c 23	20	100.0	116	3	BQ372956
c 24	20	100.0	116	6	CD521979
c 25	20	100.0	116	8	NB3678
c 26	20	100.0	118	1	AA004455
c 27	20	100.0	118	8	H57636
c 28	20	100.0	120	1	AA007673
c 29	20	100.0	120	1	AA013351
c 30	20	100.0	120	1	AA367235
c 31	20	100.0	122	2	BG956764
c 32	20	100.0	122	2	BG957876
c 33	20	100.0	122	2	BG957892
c 34	20	100.0	123	7	CK854800
c 35	20	100.0	123	8	R92902
c 36	20	100.0	125	2	BF820510
c 37	20	100.0	125	8	H71447
c 38	20	100.0	125	8	H73044
c 39	20	100.0	125	8	H73847
c 40	20	100.0	126	1	AA579709
c 41	20	100.0	126	9	AQ201947
c 42	20	100.0	128	1	AA328830
c 43	20	100.0	128	8	H67228
c 44	20	100.0	128	8	T93695
c 45	20	100.0	130	1	AA659815

ALIGNMENTS

RESULT 1
AA054107/c
LOCUS
DEFINITION
IMAGE:380387 5', similar to gb:X77738_rnal BAND 3 ANION TRANSPORT PROTEIN (HUMAN) ; , mRNA sequence.
ACCESSION
AA054107
VERSION
AA054107.1 GI:1545030
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
REFERENCE
Hominidae; Homo.
AUTHORS
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlff,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.
TITLE
The WashU-Merck EST Project
JOURNAL
Unpublished (1995)
COMMENT
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1..38
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:1288644"
/db_xref="taxon:9606"
/clone="IMAGE:380387"
/sex="male"
/tissue type="retina"
/dev stage="55 year old"
/lab_hosts="DH10B (ampicillin resistant)"
/clone_lib="Soares retina N2b4HR"

/note="Organ: eye; Vector: pTT3D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTACCAATCTGAAGTGGAGCGCCGCTTTTCTTTTCTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pTT3 vector (Pharmacia). The retinas were obtained from a 55 year old Caucasian and total cellular poly(A)+ RNA was extracted 6 hrs after their removal. The retina RNA was kindly provided by Roderick R. McInnes M.D. Ph.D. from the University of Toronto. Library constructed by Bento Soares and M.Fatima Bonaldo. "

ORIGIN

Query Match 100.0%; Score 20; DB 1; Length 38;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTGCTCTGTGGCCAGGCT 20
 |||||
 Db 28 CTTGCTCTGTGGCCAGGCT 9

RESULT 2

AA911358
 LOCUS
 DEFINITION
 o676d10.81 NCI CGAP Lu5 Homo sapiens cDNA clone IMAGE:1417555 3' similar to gb:X01057_rnal INTERLEUKIN-2 RECEPTOR ALPHA CHAIN PRECURSOR (HUMAN); contains element MER22 repetitive element ;, mRNA sequence.

ACCESSION

AA911358.1 GI:3050722

VERSION

AA911358

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 39)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgaps-rc@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio1.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality

Insert Length: 454 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 1.

Location/Qualifiers

1. .39

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:1417555"

/tissue_type="carcinoid"

/lab_host="DH10B"

/clone_lib="NCI CGAP Lu5"

/note="Organ: lung; Vector: pTT3D-Pac (Pharmacia) with a

modified polylinker; 1st strand cDNA was prepared from

neuroendocrine lung carcinoid, and was then primed with a

Not I - oligo(dT) primer. Double-stranded cDNA was ligated

to Eco RI adaptors (Pharmacia), digested with Not I and

ORIGIN

Query Match 100.0%; Score 20; DB 1; Length 39;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTGCTCTGTGGCCAGGCT 20
 |||||
 Db 8 CTTGCTCTGTGGCCAGGCT 27

RESULT 3

AZ576121

LOCUS

DEFINITION

sapiens genomic 5', genomic survey sequence.

AZ576121

AZ576121.1 GI:11562432

GSS.

KEYWORDS

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 46)

Henkel, G., Liyanage, M., Pratt, E., Huang, D., Riley, M.,

Bernardino, A., Durick, K. and Pollok, B.

Exon-trap tags from a T47D GenomeScreen(TM) Library

Unpublished (2000)

Contact: Greg Henkel

Gene Expression

Aurora Biosciences Corp.

11010 Torreyana Road, San Diego, CA 92121, USA

Tel: 8584048436

Fax: 8584046719

Email: henkelg@aurorabio.com

Pools of cells were isolated from a GenomeScreen(TM) library. The

library of cells was generated by retroviral integration of a gene

tagging element consisting of: 1) A promoterless beta-lactamase

preceded by a splice acceptor as a reporter for gene expression;

2) A promoter driving neomycin resistance followed by a splice

donor to trap downstream exons. 3' RACE from neomycin gene was

performed using total RNA from isolated pools. Output was shotgun

cloned in pAMP-1 and used to transform DH5-alpha competent

bacteria. 5' ends of reported sequences were immediately preceded

by splice donor from the trapping construct.

Class: exon-trapped.

Location/Qualifiers

1. .46

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/tissue_type="Carcinoma"

/cell_type="Epithelial"

/cell_line="T47D"

/clone_lib="Genetrap T47D Human Breast Carcinoma Library"

/notes="Organ: Breast; Vector: pAMP-1; 3' RACE of total RNA

from genetrap pools; shotgun clone in pAMP-1 and used to

transform DH5-alpha competent bacteria."

ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 46;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTGCTCTGTGGCCAGGCT 20
 |||||
 Db 6 CTTGCTCTGTGGCCAGGCT 25

cloned into the Not I and Eco RI sites of the modified pTT3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo. "

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 14:15:14 ; Search time 271.36 Seconds
(without alignments)
491.207 Million cell updates/sec

Title: US-10-754-446-3

Perfect score: 20

Sequence: 1 cttgctctgttgcaggct 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 21.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*
- 14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	14	AEB28959 Human MCO
c 2	20	100.0	24	6	ABV75668 Human PD1
3	20	100.0	41	4	AAH48082 Ras GRPas
c 4	20	100.0	41	6	AAE15951 Human pro
c 5	20	100.0	41	6	AAE15950 Human pro
6	20	100.0	41	10	ADC59511 Polypepti
c 7	20	100.0	51	4	AAI79698 Human con
c 8	20	100.0	51	4	AAH89487 Human cod
c 9	20	100.0	51	4	AAH89486 Human cod
10	20	100.0	60	10	ADI20565 Oligonucle
11	20	100.0	71	8	ABZ09117 Human oli
12	20	100.0	71	10	ABZ78570 Tumour su
c 13	20	100.0	77	8	ABZ09693 Human oli
c 14	20	100.0	77	10	ABZ79146 Tumour su
15	20	100.0	82	13	ADR16680 MicroRNA
16	20	100.0	84	4	AAE28538 Genomic s
17	20	100.0	84	10	ADGA1734 Human res
18	20	100.0	84	11	ADI97508 Human res
c 19	20	100.0	91	2	AAE26410 Human gen

c 20	20	100.0	93	4	AAK65580 Human imm
c 21	20	100.0	94	4	AAK73464 Human imm
c 22	20	100.0	94	4	AAK72995 Human imm
23	20	100.0	95	4	AAK77204 Human imm
24	20	100.0	95	4	AAK77203 Human imm
c 25	20	100.0	95	4	AAK70695 Human imm
c 26	20	100.0	98	3	AAK15933 Human sec
c 27	20	100.0	98	4	AAE36698 Human car
c 28	20	100.0	98	4	AAK71993 Human imm
c 29	20	100.0	98	10	AAE47392 Human car
c 30	20	100.0	98	13	ADJ08810 Human car
31	20	100.0	102	3	AAK13267 Human sec
c 32	20	100.0	102	4	AAE36278 Human car
c 33	20	100.0	102	10	AAE46972 Human car
c 34	20	100.0	102	13	ADJ08390 Human car
c 35	20	100.0	103	4	AAE28603 Genomic s
c 36	20	100.0	103	10	ADGA1799 Human res
c 37	20	100.0	103	11	ADI97573 Human res
c 38	20	100.0	104	4	AAE107415 Human rep
c 39	20	100.0	104	5	AAE40769 DNA encod
c 40	20	100.0	104	11	ADJ09975 Human pro
c 41	20	100.0	106	12	ADO57414 DNA encod
c 42	20	100.0	111	4	AAK74404 Human imm
c 43	20	100.0	113	5	ABAE13542 Human ner
c 44	20	100.0	114	3	AAE21248 Human sec
c 45	20	100.0	114	4	AAK77447 Human imm

ALIGNMENTS

RESULT 1

AEB28959
ID AEB28959 standard; DNA; 20 BP.

AC AEB28959;

XX 22-SEP-2005 (first entry)

XX Human MCOLN1 gene PCR primer PRI F1 SEQ ID NO:3.

XX mucolipidosis IV; high throughput screening; mucolipin-1; MCOLN1; PCR; primer; ss.

OS Homo sapiens.

XX US2005153300-A1.

PD 14-JUL-2005.

XX 09-JAN-2004; 2004US-00754446.

XX 09-JAN-2004; 2004US-00754446.

XX (QUES-) QUEST DIAGNOSTICS INC.

XX Sun W, Hantash F;

XX WPI; 2005-521160/53.

XX Diagnosing mucolipidosis (ML) IV mutant sequence in nucleic acids, by amplifying the nucleic acid, detecting amplified product with labeled oligonucleotide probes via a change in fluorescence which indicates the presence of an ML IV mutant.

XX Claim 1; SEQ ID NO 3; 15pp; English.

XX The invention relates to a method (M1) for detecting the presence of mucolipidosis IV (ML IV) mutant sequences in a nucleic acid. (M1) involves contacting the nucleic acid with oligonucleotide primers and probes, conducting amplification by temperature cycling and monitoring the accumulation of amplified nucleic acid by detecting an increase in donor fluorophore fluorescence or decrease in acceptor fluorophore

CC fluorescence, which indicates the presence of ML IV mutant sequence. Also
 CC described: (1) an oligonucleotide (ON1) comprising a sequence
 CC complementary to the coding or non-coding strand of the mucolin-1
 CC (MCOLN1) gene; and (2) a kit (K1) for amplifying sequences of ML IV
 CC sequences, comprising ON1. (M1) is useful for determining the presence of
 CC ML IV mutant sequence in a nucleic acid, and/or wild-type sequences. (M1)
 CC is useful for genetic counseling of individuals at risk for ML IV or at
 CC risk to having offspring afflicted with ML IV. The MCOLN1 gene is located
 CC on chromosome 19p13. The present sequence represents a PCR primer for the
 CC MCOLN1 gene, which is used in the exemplification of the present
 CC invention.

XX
 SQ Sequence 20 BP; 1 A; 7 C; 5 G; 7 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 14; Length 20;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTGCTCTGTGCCAGGCT 20
 |||||
 Db 1 CTTGCTCTGTGCCAGGCT 20

RESULT 2
 ABV75668/c
 ID ABV75668 standard; DNA; 24 BP.
 XX
 AC ABV75668;
 XX
 DT 28-JAN-2003 (first entry)
 XX
 DE Human FD16.06 PCR primer 2.
 XX
 KW Human; FD16.06; cancer; HIV; PCR; primer; ss.
 XX
 OS Homo sapiens.
 XX
 PN CN1351034-A.
 XX
 PD 29-MAY-2002.
 XX
 PF 26-OCT-2000; 2000CN-00125785.
 XX
 PR 26-OCT-2000; 2000CN-00125785.
 XX
 PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.
 XX
 PI Mao Y, Xie Y;
 XX
 DR WPI; 2002-619857/67.
 XX
 PT Polypeptide-human FD16.06 and polynucleotide for coding it, useful for
 XX treating diseases such as cancer and HIV infection.

XX Example 2; Page 16 (Disclosure); 33pp; Chinese.
 XX The invention relates to a novel polypeptide, human FD16.06, and the
 XX polynucleotide encoding it. The antagonist against the polypeptide is
 XX also disclosed. The polypeptide is useful for treating diseases such as
 XX cancer and HIV infection. The present sequence represents a PCR primer
 XX used to amplify the human FD16.06 gene of the invention

XX
 SQ Sequence 24 BP; 9 A; 7 C; 7 G; 1 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 24;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTGCTCTGTGCCAGGCT 20
 |||||
 Db 22 CTTGCTCTGTGCCAGGCT 3

RESULT 3
 AAH48082
 ID AAH48082 standard; DNA; 41 BP.
 XX
 AC AAH48082;
 XX
 DT 19-SEP-2001 (first entry)
 XX
 DE Ras GTPase-activating protein 12 probe #2.
 XX

Ras GTPase-activating protein 12; cytostatic; Antiinflammatory;
 KW Immunomodulatory; Anti-HIV; malignant tumour; haemopathy; HIV infection;
 KW immunological disease; inflammation; gene therapy; probe; ss.

XX Unidentified.

XX WO200147998-A1.
 XX 05-JUL-2001.

XX 25-DEC-2000; 2000WO-CN000680.

XX 27-DEC-1999; 99CN-00125367.

XX (SHAN-) SHANGHAI BIOWINDOW GENE DEV INC.

XX Mao Y, Xie Y;

XX WPI; 2001-418231/44.

XX RAS GTPase-activating protein 12 and encoded polynucleotide, applicable
 PT in diagnosis and treatment of malignant tumor, hemopathy, HIV infection,
 PT immunological diseases and various inflammation.

XX Example 7; Page 20; 38pp; Chinese.

XX The present invention relates to Ras GTPase-activating protein 12 and its
 CC coding sequence (see AAH48076 and AAG64227). The GTPase-activating
 CC protein and its coding sequence are useful in the diagnosis and treatment
 CC of malignant tumour, haemopathy, HIV infection, immunological diseases
 CC and various inflammations. The present sequence is a probe, which was
 CC used in an example from the present invention

XX SQ Sequence 41 BP; 6 A; 10 C; 12 G; 13 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 4; Length 41;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTGCTCTGTGCCAGGCT 20
 |||||
 Db 13 CTTGCTCTGTGCCAGGCT 32

RESULT 4

AAS15951/c

ID AAS15951 standard; DNA; 41 BP.

XX AAS15951;

XX 14-FEB-2002 (first entry)

XX Human proteolytic enzyme regulatory protein 11, probe #2.

XX Human; proteolytic enzyme regulatory protein 11; cytostatic; virucidal;
 KW immunomodulatory; antiinflammatory; haemostatic; cancer; haemopathy;
 KW human immunodeficiency virus; HIV; immunological disease; inflammation;
 KW embryonic development disorder; developmental disorder; probe; ss.

XX Homo sapiens.

XX WO200174867-A1.

XX

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 16:41:33 ; Search time 483.04 Seconds
(without alignments)
342.389 Million cell updates/sec

Title: US-10-754-446-3

Perfect score: 20

Sequence: 1 CTTGCTCTGTGCCAGGCT 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	9	US-10-754-446-3
2	20	100.0	60	8	US-10-804-859-55
3	20	100.0	71	9	US-10-467-851-277
4	20	100.0	72	9	US-10-708-204-6674
5	20	100.0	77	9	US-10-467-851-853
6	20	100.0	78	9	US-10-708-204-6268
7	20	100.0	78	9	US-10-708-204-6998
8	20	100.0	83	9	US-10-708-204-6598
9	20	100.0	84	3	US-09-764-860-972
10	20	100.0	84	5	US-10-074-035-972
11	20	100.0	84	6	US-10-212-872-972
12	20	100.0	84	9	US-10-708-204-6304
13	20	100.0	86	9	US-10-708-204-5760
14	20	100.0	87	9	US-10-708-204-6688
15	20	100.0	88	9	US-10-708-204-6626
16	20	100.0	88	9	US-10-708-204-6676
17	20	100.0	88	9	US-10-708-204-6676
18	20	100.0	88	9	US-10-708-204-6676
19	20	100.0	88	9	US-10-708-204-6699
20	20	100.0	88	9	US-10-708-204-6699
21	20	100.0	88	9	US-10-708-204-6784
22	20	100.0	90	9	US-10-708-204-6710
23	20	100.0	92	9	US-10-708-204-6788

ALIGNMENTS

RESULT 1
US-10-754-446-3
; Sequence 3, Application US/10754446
; Publication No. US20050153300A1
; GENERAL INFORMATION:
; APPLICANT: SUN, WEIMIN
; APPLICANT: HANTASH, PERAS
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DETECTION OF
; TITLE OF INVENTION: MUCOLIPIDOSIS IV MUTATIONS
; FILE REFERENCE: 034827-2301
; CURRENT APPLICATION NUMBER: US/10/754,446
; CURRENT FILING DATE: 2004-01-09
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 3.2
; SEQ ID NO 3
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-10-754-446-3

Query Match 100.0%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTGCTCTGTGCCAGGCT 20
|||||
Db 1 CTTGCTCTGTGCCAGGCT 20

RESULT 2
US-10-804-859-55
; Sequence 55, Application US/10804859
; Publication No. US20040265865A1
; GENERAL INFORMATION:
; APPLICANT: The University of Queensland
; TITLE OF INVENTION: A method for identifying effector molecules for gene network inte
; FILE REFERENCE: 2563972/EJH
; CURRENT APPLICATION NUMBER: US/10/804,859
; CURRENT FILING DATE: 2004-03-19
; PRIOR APPLICATION NUMBER: US 60/324127
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 55

Sequence 6492, Ap
Sequence 6445, Ap
Sequence 6753, Ap
Sequence 6848, Ap
Sequence 5730, Ap
Sequence 6816, Ap
Sequence 6419, Ap
Sequence 6198, Ap
Sequence 2198, Ap
Sequence 2198, Ap
Sequence 6436, Ap
Sequence 5772, Ap
Sequence 6371, Ap
Sequence 6374, Ap
Sequence 6374, Ap
Sequence 7165, Ap
Sequence 1778, Ap
Sequence 1778, Ap
Sequence 1778, Ap
Sequence 1037, Ap
Sequence 1037, Ap
Sequence 1037, Ap

; LENGTH: 60
; TYPE: DNA
; ORGANISM: human
US-10-804-859-55

Query Match 100.0%; Score 20; DB 8; Length 60;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTGCTCTGTGCCAGGCT 20
|||||
DB 7 CTTGCTCTGTGCCAGGCT 26

RESULT 3

US-10-467-851-277
; Sequence 277, Application US/10467851
; Publication No. US20050221303A1
; GENERAL INFORMATION:
; APPLICANT: TELERMAN, Adam
; APPLICANT: AMSON, Robert
; APPLICANT: TUIJNDER, Marcel,
; APPLICANT: SUSINI, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; TITLE OF INVENTION: AND THEIR USE AS MEDICINES
; FILE REFERENCE: 11416-014-999
; CURRENT APPLICATION NUMBER: US/10/467,851
; CURRENT FILING DATE: 2003-08-13
; PRIOR FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: PCT/FR 02/00 543
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: FR 01/01 925
; NUMBER OF SEQ ID NOS: 1020
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 277
; LENGTH: 71
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-10-467-851-277

Query Match 100.0%; Score 20; DB 9; Length 71;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTGCTCTGTGCCAGGCT 20
|||||
DB 11 CTTGCTCTGTGCCAGGCT 30

RESULT 4

US-10-708-204-6674/c
; Sequence 6674, Application US/10708204
; Publication No. US20050222399A1
; GENERAL INFORMATION:
; APPLICANT: ROSETTA GENOMICS LTD
; TITLE OF INVENTION: BIOINFORMATICALLY DETECTABLE GROUP OF NOVEL REGULATORY
; TITLE OF INVENTION: OLIGONUCLEOTIDES ASSOCIATED WITH ALZHEIMER'S DISEASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 55033
; CURRENT APPLICATION NUMBER: US/10/708,204
; CURRENT FILING DATE: 2004-02-16
; NUMBER OF SEQ ID NOS: 7351
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 6674
; LENGTH: 72
; TYPE: RNA
; ORGANISM: Homo Sapiens
US-10-708-204-6674

Query Match 100.0%; Score 20; DB 9; Length 72;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTGCTCTGTGCCAGGCT 20
|||||
DB 67 CTTGCTCTGTGCCAGGCT 48

RESULT 5

US-10-467-851-853/c
; Sequence 853, Application US/10467851
; Publication No. US20050221303A1
; GENERAL INFORMATION:
; APPLICANT: TELERMAN, Adam
; APPLICANT: AMSON, Robert
; APPLICANT: TUIJNDER, Marcel,
; APPLICANT: SUSINI, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; TITLE OF INVENTION: AND THEIR USE AS MEDICINES
; FILE REFERENCE: 11416-014-999
; CURRENT APPLICATION NUMBER: US/10/467,851
; CURRENT FILING DATE: 2003-08-13
; PRIOR FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: PCT/FR 02/00 543
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: FR 01/01 925
; NUMBER OF SEQ ID NOS: 1020
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 853
; LENGTH: 77
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-10-467-851-853

Query Match 100.0%; Score 20; DB 9; Length 77;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTGCTCTGTGCCAGGCT 20
|||||
DB 39 CTTGCTCTGTGCCAGGCT 20

RESULT 6

US-10-708-204-6268/c
; Sequence 6268, Application US/10708204
; Publication No. US20050222399A1
; GENERAL INFORMATION:
; APPLICANT: ROSETTA GENOMICS LTD
; TITLE OF INVENTION: BIOINFORMATICALLY DETECTABLE GROUP OF NOVEL REGULATORY
; TITLE OF INVENTION: OLIGONUCLEOTIDES ASSOCIATED WITH ALZHEIMER'S DISEASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 55033
; CURRENT APPLICATION NUMBER: US/10/708,204
; CURRENT FILING DATE: 2004-02-16
; NUMBER OF SEQ ID NOS: 7351
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 6268
; LENGTH: 78
; TYPE: RNA
; ORGANISM: Homo Sapiens
US-10-708-204-6268

Query Match 100.0%; Score 20; DB 9; Length 78;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTGCTCTGTGCCAGGCT 20
|||||
DB 78 CTTGCTCTGTGCCAGGCT 59

RESULT 7

US-10-708-204-6998/c

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 16:32:47 ; Search time 72.8 Seconds
(without alignments)
488.341 Million cell updates/sec

Title: US-10-754-446-3

Perfect score: 20

Sequence: 1 cttgctctgtgcccaggct 20

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCUS COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RB COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	20	100.0	98	3	US-09-513-999C-20008
c 2	20	100.0	102	3	US-09-513-999C-17342
c 3	20	100.0	114	3	US-09-513-999C-25323
c 4	20	100.0	117	3	US-09-513-999C-17724
c 5	20	100.0	117	3	US-09-513-999C-24469
c 6	20	100.0	128	3	US-09-513-999C-15044
c 7	20	100.0	145	3	US-09-621-976-14390
c 8	20	100.0	145	3	US-09-513-999C-16178
c 9	20	100.0	145	3	US-09-513-999C-25910
c 10	20	100.0	151	3	US-09-513-999C-24127
c 11	20	100.0	152	3	US-09-621-976-11650
c 12	20	100.0	152	3	US-09-621-976-13583
c 13	20	100.0	155	3	US-09-513-999C-29581
c 14	20	100.0	158	3	US-09-621-976-11688
c 15	20	100.0	158	3	US-09-621-976-12407
c 16	20	100.0	166	3	US-09-621-976-11848
c 17	20	100.0	168	3	US-09-621-976-11602
c 18	20	100.0	170	3	US-09-621-976-11644
c 19	20	100.0	170	3	US-09-621-976-11777
c 20	20	100.0	170	3	US-09-621-976-11869
c 21	20	100.0	170	3	US-09-621-976-12011
c 22	20	100.0	170	3	US-09-621-976-12076
c 23	20	100.0	170	3	US-09-621-976-12098
c 24	20	100.0	170	3	US-09-621-976-12260

ALIGNMENTS

RESULT 1

US-09-513-999C-20008/c
; Sequence 20008, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513.999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 20008
; LENGTH: 98
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-20008

Query Match 100.0%; Score 20; DB 3; Length 98;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTGCTCTGTGCCCAGGCT 20
|||||
Db 69 CTTGCTCTGTGCCCAGGCT 50

RESULT 2

US-09-513-999C-17342
; Sequence 17342, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513.999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26

Sequence 12300, A
Sequence 12391, A
Sequence 13487, A
Sequence 13529, A
Sequence 13550, A
Sequence 11724, A
Sequence 11808, A
Sequence 11926, A
Sequence 12242, A
Sequence 13295, A
Sequence 11573, A
Sequence 11788, A
Sequence 11924, A
Sequence 12171, A
Sequence 12205, A
Sequence 12233, A
Sequence 11638, A
Sequence 11713, A
Sequence 12010, A
Sequence 12151, A

; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 17342
; LENGTH: 102
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-17342

Query Match 100.0%; Score 20; DB 3; Length 102;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTGCTCTGTGCCAGGCT 20
|||||
Db 78 CTTGCTCTGTGCCAGGCT 97

RESULT 3

US-09-513-999C-25323/c
; Sequence 25323, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 25323
; LENGTH: 114
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-25323

Query Match 100.0%; Score 20; DB 3; Length 114;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTGCTCTGTGCCAGGCT 20
|||||
Db 91 CTTGCTCTGTGCCAGGCT 72

RESULT 4

US-09-513-999C-17724
; Sequence 17724, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 17724
; LENGTH: 117
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9

; OTHER INFORMATION: k=g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11
; OTHER INFORMATION: k=g or t
US-09-513-999C-17724

Query Match 100.0%; Score 20; DB 3; Length 117;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTGCTCTGTGCCAGGCT 20
|||||
Db 25 CTTGCTCTGTGCCAGGCT 44

RESULT 5

US-09-513-999C-24469/c
; Sequence 24469, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 24469
; LENGTH: 117
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-24469

Query Match 100.0%; Score 20; DB 3; Length 117;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTGCTCTGTGCCAGGCT 20
|||||
Db 112 CTTGCTCTGTGCCAGGCT 93

RESULT 6

US-09-513-999C-15044
; Sequence 15044, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 15044
; LENGTH: 128
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-15044

Query Match 100.0%; Score 20; DB 3; Length 128;
Best Local Similarity 100.0%; Pred. No. 3.7;

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 15:49:05 ; Search time 1080.46 Seconds
(without alignments)
1157.425 Million cell updates/sec

Title: US-10-754-446-2

Perfect score: 22

Sequence: 1 taaccaccatcggaatcatgtc 22

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_in.*

3: gb_env.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pr.*

9: gb_ro.*

10: gb_sta.*

11: gb_sy.*

12: gb_un.*

13: gb_vi.*

14: gb_htg.*

15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	22	100.0	790	8 AF305572S3	AF305574 Homo sapi
C 2	22	100.0	1740	6 AX280021	AX280021 Sequence
C 3	22	100.0	1829	6 C0719994	C0719994 Sequence
C 4	22	100.0	2004	8 AF249319	AF249319 Homo sapi
C 5	22	100.0	2037	8 HSA233970	HSA233970 Homo sapi
C 6	22	100.0	2049	8 AK026102	AK026102 Homo sapi
C 7	22	100.0	2051	8 AF287269	AF287269 Homo sapi
C 8	22	100.0	2052	6 AX083508	AX083508 Sequence
C 9	22	100.0	2063	8 AK222673	AK222673 Homo sapi
C 10	22	100.0	2078	8 BC005149	BC005149 Homo sapi
C 11	22	100.0	2094	6 BD233734	BD233734 31 human
C 12	22	100.0	2095	6 AX280019	AX280019 Sequence
C 13	22	100.0	2272	8 HSA293659	HSA293659 Homo sapi
C 14	22	100.0	13270	8 AF287270	AF287270 Homo sapi
C 15	22	100.0	155645	14 AC021153	AC021153 Homo sapi
C 16	22	100.0	173126	8 AC008878	AC008878 Homo sapi
C 17	20.4	92.7	2065	8 AB125179	AB125179 Macaca fa
C 18	20	90.9	110000	1 AB014295_10	Continuation (11 o

C 19	20	90.9	349980	6 AX492785	AX492785 Sequence
C 20	20	90.9	349980	6 AX553952	AX553952 Sequence
C 21	17.8	80.9	717	10 BV552653	BV552653 S221P6043
C 22	17.8	80.9	734	10 BV539631	BV539631 G591P6424
C 23	17.8	80.9	767	10 BV534108	BV534108 G591P6224
C 24	17.8	80.9	775	10 BV539839	BV539839 G591P6408
C 25	17.8	80.9	826	10 BV532664	BV532664 G591P6209
C 26	17.8	80.9	1347	13 DQ103713	DQ103713 Influenza
C 27	17.8	80.9	3666	1 AB050784	AB050784 Spingomo
C 28	17.8	80.9	7722	1 AB032203	AB032203 Spingomo
C 29	17.8	80.9	24756	14 AC010403	AC010403 Homo sapi
C 30	17.8	80.9	69678	8 AC091573	AC091573 Homo sapi
C 31	17.8	80.9	143803	5 BX470131	BX470131 Zebrafish
C 32	17.8	80.9	157076	8 AC019031	AC019031 Homo sapi
C 33	17.8	80.9	174578	8 AC013287	AC013287 Homo sapi
C 34	17.8	80.9	177301	8 AC078868	AC078868 Homo sapi
C 35	17.8	80.9	182182	5 BX890617	BX890617 Zebrafish
C 36	17.8	80.9	212671	14 AC084849	AC084849 Homo sapi
C 37	17.8	80.9	213611	8 AF276759	AF276759 Homo sapi
C 38	17.8	80.9	242704	8 BS000230	BS000230 Pan trogl
C 39	17.4	79.1	3153	6 E15417	E15417 Arthropode
C 40	17.4	79.1	3700	1 STHISX	V01373 Salmonella
C 41	17.4	79.1	4483	1 STYHISTO	J01805 S. typhimur
C 42	17.4	79.1	8516	1 D88362	D88362 Arthropode
C 43	17.4	79.1	20518	1 AE008805	AE008805 Salmonell
C 44	17.4	79.1	40699	2 CEZK637	Z1115 Caenorhabdi
C 45	17.4	79.1	82568	15 AP004906	AP004906 Lotus cor

ALIGNMENTS

RESULT 1
AF305572S3/c
LOCUS AF305572S3 790 bp DNA linear PRI 26-DEC-2000
DEFINITION Homo sapiens mucopolipin 1 (MCOLN1) gene, exons 3 and 4.
ACCESSION AF305574
VERSION AF305574.1 GI:11991200

SEGMENT 3 of 8

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE 1 (bases 1 to 790)

AUTHORS Bargal, R., Avidan, N., Ben-Asher, E., Olender, Z., Zeigler, M.,

Frumkin, A., Raas-Rothschild, A., Glusman, G., Lancet, D. and Bach, G.

TITLE Identification of the gene causing mucopolipidosis type IV

JOURNAL Nat. Genet. 26 (1), 118-123 (2000)

PUBMED 10973263

REFERENCE 2 (bases 1 to 790)

AUTHORS Bargal, R., Avidan, N., Ben-Asher, E., Olender, A., Zeigler, M.,

Frumkin, A., Raas-Rothschild, A., Glusman, G., Lancet, D. and Bach, G.

TITLE Direct Submission

JOURNAL Submitted (13-SEP-2000) Molecular Genetics, The Weizmann Institute

of Science, P. O. Box 26, Rehovot 76100, Israel

FEATURES Location/Qualifiers

source 1..790

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

267..434

/gene="MCOLN1"

/number=3

589..757

/gene="MCOLN1"

/number=4

ORIGIN

Query Match 100.0%; Score 22; DB 8; Length 790;

Best Local Similarity 100.0%; Pred. No. 0.87; Indels 0; Gaps 0;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACCACCATCGGATCAATGTC 22
Db 750 TAACCACCATCGGATCAATGTC 729

RESULT 2
AX280021/c
LOCUS AX280021 1740 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 3 from Patent WO0177331.
ACCESSION AX280021
VERSION AX280021.1 GI:16607475
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1
AUTHORS Curtis, R.A. and Silos-Santiago, I.
TITLE Human tripartite channel protein-2 (tlcc-2)
JOURNAL Patent: WO 0177331-A 3 18-OCT-2001;
MILLENNIUM PHARMACEUTICALS, INC. (US)

FEATURES
source
1. .1740
/location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1. .>1740
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD10547.1"
/db_xref="GI:16607476"
/translation="MTAPAGPRGSETERLLTPNGYGTQAGSPAPPTPEEDLRRL
LKTFSPCKFRAGKPKMLQVVKILVTVQLFGLSNQLAVTFREENTIAFR
HLFLGYSQADDTFAAYTREQLYQAIHFAVDQYALPDLVSLGRYAYVGGGDPWTNG
SGLALCQRYHGHVDPANDTFIDPMVVTDCIQVDPPEPPPPSDLLTLLSSSY
KNLTKFHLVNTIHFRLKTNLQSLINNEIPDCYFVSLITFDNKAHSGRIPISLE
TOAHIOECKHPSPVQHGDSFRLLFDVVVILTCSLFCLCARSLRGLLQNEFVGFM
WRQGRVISLWERLEFVNGWYLLVTSVLTSGTINKIGIEAKNLASYDVCSILIGT
HVLVWGVIYRLITFFHNYNLIATLRVALPSPVNRCCCVAVIYLGVCFCGWLVLGY
STKFRSLWSVSECLSLNGDDMPVTFPAAMQAGRSSLWFLFSQLYLSFISLFYIM
VLSLFTALITGAYDTIKHPGAGAESELOQYIAQCDSPTSGKFRGSGSACSLCC
CGRDPSEHSLVN"

CDS
1. .1740
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD10547.1"
/db_xref="GI:16607476"
/translation="MTAPAGPRGSETERLLTPNGYGTQAGSPAPPTPEEDLRRL
LKTFSPCKFRAGKPKMLQVVKILVTVQLFGLSNQLAVTFREENTIAFR
HLFLGYSQADDTFAAYTREQLYQAIHFAVDQYALPDLVSLGRYAYVGGGDPWTNG
SGLALCQRYHGHVDPANDTFIDPMVVTDCIQVDPPEPPPPSDLLTLLSSSY
KNLTKFHLVNTIHFRLKTNLQSLINNEIPDCYFVSLITFDNKAHSGRIPISLE
TOAHIOECKHPSPVQHGDSFRLLFDVVVILTCSLFCLCARSLRGLLQNEFVGFM
WRQGRVISLWERLEFVNGWYLLVTSVLTSGTINKIGIEAKNLASYDVCSILIGT
HVLVWGVIYRLITFFHNYNLIATLRVALPSPVNRCCCVAVIYLGVCFCGWLVLGY
STKFRSLWSVSECLSLNGDDMPVTFPAAMQAGRSSLWFLFSQLYLSFISLFYIM
VLSLFTALITGAYDTIKHPGAGAESELOQYIAQCDSPTSGKFRGSGSACSLCC
CGRDPSEHSLVN"

Qy 1 TAACCACCATCGGATCAATGTC 22
Db 568 TAACCACCATCGGATCAATGTC 547

RESULT 3
CQ719994/c
LOCUS CQ719994 1829 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 5928 from Patent WO02068579.
ACCESSION CQ719994
VERSION CQ719994.1 GI:42280851
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
human exons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 5928 06-SEP-2002;

PE Corporation (NY) (US)
Location/Qualifiers
1. .1829
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 22; DB 6; Length 1829;
Best Local Similarity 100.0%; Pred. No. 0.96; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACCACCATCGGATCAATGTC 22
Db 537 TAACCACCATCGGATCAATGTC 516

RESULT 4
AF249319/c
LOCUS AF249319 2004 bp mRNA linear PRI 22-SEP-2000
DEFINITION Homo sapiens mucopolipidosis type IV protein (MLIV) mRNA, complete
cds.
ACCESSION AF249319
VERSION AF249319.1 GI:9971787
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2004)
AUTHORS Bargal, R., Avidan, N., Ben-Asher, E., Olender, Z., Zeigler, M.,
Frumkin, A., Rases-Rothschild, A., Glusman, G., Lancet, D. and Bach, G.
TITLE Identification of the gene causing mucopolipidosis type IV
JOURNAL Nat. Genet. 26 (1), 118-123 (2000)
PUBMED 10973263
REFERENCE 2 (bases 1 to 2004)
AUTHORS Bargal, R., Avidan, N., Ben-Asher, E., Olender, Z., Zeigler, M.,
Frumkin, A., Rases-Rothschild, A., Glusman, G., Lancet, D. and Bach, G.
TITLE Direct Submission
JOURNAL Submitted (26-NOV-2000) Molecular Genetics, The Weizmann Institute
of Science, P. O. Box 26, Rehovot 76100, Israel

FEATURES
source
1. .2004
/location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. .2004
/genes="MLIV"
85. .1827
/genes="MLIV"
/note="causes mucopolipidosis type IV"
/codon_start=1
/product="mucopolipidosis type IV protein"
/protein_id="AAG10422.1"
/db_xref="GI:9971788"

ORIGIN
Query Match 100.0%; Score 22; DB 8; Length 2004;
Best Local Similarity 100.0%; Pred. No. 0.97; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 15:55:50 ; Search time 2460.3 Seconds
(without alignments)
418.369 Million cell updates/sec

Title: US-10-754-446-2

Perfect score: 22

Sequence: 1 taaccaccatcgatcaatgctc 22

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_hic.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_est7.*
9: gb_gss1.*
10: gb_gss2.*
11: gb_gss3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	22	100.0	423	1	AW969529 EST381606
C 2	22	100.0	425	3	BP391004 BP391004
C 3	22	100.0	439	8	T66341 yc78b06.t1
C 4	22	100.0	480	1	AW328429 ds02e06.x
C 5	22	100.0	510	2	BG272754 nah89c03
C 6	22	100.0	554	3	BP265945 BP265945
C 7	22	100.0	580	3	BP266237 BP266237
C 8	22	100.0	581	3	BP332355 BP332355
C 9	22	100.0	657	3	B1819028 60303154
C 10	22	100.0	660	3	B1836168 603085896
C 11	22	100.0	660	3	B1907877 603066872
C 12	22	100.0	663	3	B1523736 603052025
C 13	22	100.0	663	3	B1836493 603082706
C 14	22	100.0	666	3	BM005772 603613386
C 15	22	100.0	666	6	CA391526 csl6c04.y
C 16	22	100.0	672	3	B1490022 603031806
C 17	22	100.0	678	8	DR001457 TC113094
C 18	22	100.0	683	6	CF132757 UI-HF-FQ0
C 19	22	100.0	696	2	BF529860 602042214
C 20	22	100.0	697	3	BM049487 603623420
C 21	22	100.0	705	3	B1822028 603040071
C 22	22	100.0	709	3	B1825945 603076515

C 23	22	100.0	743	3	BI771045
C 24	22	100.0	790	6	CB988285
C 25	22	100.0	810	3	BI753742
C 26	22	100.0	818	2	BG831705
C 27	22	100.0	846	5	EX383671
C 28	22	100.0	873	6	CD558014
C 29	22	100.0	890	5	EX399044
C 30	22	100.0	901	5	BUS42887
C 31	22	100.0	902	5	BQ901107
C 32	22	100.0	909	5	BUS43879
C 33	22	100.0	910	5	BUS52585
C 34	22	100.0	962	5	BX402300
C 35	22	100.0	998	3	BI824195
C 36	22	100.0	998	3	BM905707
C 37	22	100.0	1026	5	EX364018
C 38	22	100.0	1047	2	BE795087
C 39	22	100.0	1050	5	EX462609
C 40	22	100.0	1078	5	EX362786
C 41	22	100.0	1144	5	EX386808
C 42	22	100.0	1743	10	AY410283
C 43	22	100.0	1996	4	CR622331
C 44	20.4	92.7	951	5	BU195645
C 45	20	90.9	540	3	BI516259

ALIGNMENTS

RESULT 1
AW969529/c
LOCUS AW969529 423 bp mRNA linear EST 01-JUN-2000
DEFINITION EST381606 MAGE resequences, MAGK Homo sapiens cDNA, mRNA sequence.
ACCESSION AW969529
VERSION AW969529.1 GI:8159373
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 423)
Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and Quackenbush,J.
TITLE Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished (2000)
COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@tigr.org
Plate: 286
Seq primer: Forward.
Location/Qualifiers
1 . 423
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGK"
/note="Vector: pBluescriptSKm"

Query Match 100.0%; Score 22; DB 1; Length 423;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACCACCATCGGATCAATGTC 22
|||||
DB 96 TAACCACCATCGGATCAATGTC 75

```

RESULT 2
BP391004/c
LOCUS
DEFINITION BP391004 Homo sapiens pancreatic islet Homo sapiens cDNA clone linear mRNA 425 bp EST 01-FEB-2005
h-t-232-79 3', mRNA sequence.
ACCESSION BP391004
VERSION BP391004.1 GI:58414498
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
1 (bases 1 to 425)
Takeda,J., Jin,L. and Horikawa,Y.
Expression profile of mRNAs from human pancreatic islet tumors
Unpublished (2005)
CONTACT: Yukio Horikawa
Laboratory of Molecular Genetics
Institute for Molecular and Cellular Regulation, Gunma University
Showa-machi 3-39-15, Maebashi, Gunma, 371-8512, Japan
Tel: 81-27-220-8832
Fax: 81-27-220-8889
Email: yhorikawa@showa.gunma-u.ac.jp.
Location/Qualifiers
1. .425
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="h-t-232-79"
/tissue_types="pancreatic islet"
/clone_lib="Homo sapiens pancreatic islet"

ORIGIN
Query Match 100.0%; Score 22; DB 3; Length 425;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACACCATCGGATCAATGTC 22
|||||
Db 73 TAACACCATCGGATCAATGTC 52

RESULT 3
T66341/c
LOCUS
DEFINITION T66341 Soares infant brain INIB Homo sapiens cDNA clone linear mRNA 439 bp EST 07-MAR-1995
IMAGE:21977 5', mRNA sequence.
ACCESSION T66341
VERSION T66341.1 GI:675386
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
1 (bases 1 to 439)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Treviski,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: get@watson.wustl.edu
Insert Size: 2264
High quality sequence stops: 345 Source: IMAGE Consortium, LLNL This
clone is available royalty-free through LLNL ; contact the IMAGE

FEATURES
source
1. .439
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:394324"
/db_xref="taxon:9606"
/clone="IMAGE:21977"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares infant brain INIB"
/notes="Organ: whole brain; Vector: Lfamid BA; Site 1: Not
I; Site 2: Hind III; 1st strand cDNA was primed with a Not
I - oligo(dT) primer [5'
AACTGGAGAAATTCGGCGCCAGGAAATTTTTCCTTTT 3'];
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the Lfamid BA vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN
Query Match 100.0%; Score 22; DB 8; Length 439;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACACCATCGGATCAATGTC 22
|||||
Db 309 TAACACCATCGGATCAATGTC 288

RESULT 4
AW328429
LOCUS
DEFINITION AW328429.1 GI:6798925
ACCESSION AW328429
VERSION AW328429.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
1 (bases 1 to 480)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
JOURNAL
COMMENT
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Edge Biosystems
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center (NISC)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Plate: LLCW0030 row: I column: 12
Seq primer: -21M13 forward primer (ABI).
Location/Qualifiers
1. .480
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2847443"
/cell_line="MGCI"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 4"
/notes="Organ: Cervix; Vector: pOTB7a; Site_1: Sclt;

Consortium (info@image.llnl.gov) for further information.
Insert Length: 2264 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 345.
Location/Qualifiers
source
1. .439
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:394324"
/db_xref="taxon:9606"
/clone="IMAGE:21977"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares infant brain INIB"
/notes="Organ: whole brain; Vector: Lfamid BA; Site 1: Not
I; Site 2: Hind III; 1st strand cDNA was primed with a Not
I - oligo(dT) primer [5'
AACTGGAGAAATTCGGCGCCAGGAAATTTTTCCTTTT 3'];
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the Lfamid BA vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 14:15:14 ; Search time 298.496 Seconds
(without alignments)
491.207 Million cell updates/sec

Title: US-10-754-446-2

Perfect score: 22

Sequence: 1 taaccaccatcgatcaatgctc 22

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

14: Geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	22	14 AEB28958	Aeb28958 Human MCO
2	22	100.0	1619	4 AAI61022	Aai61022 Human pol
3	22	100.0	1740	6 AAI17100	Aai17100 Human TRP
4	22	100.0	1740	6 ABL40755	Abl40755 Human TLC
5	22	100.0	1741	4 AAI59236	Aai59236 Human pol
6	22	100.0	1743	9 ADB84284	Adb84284 Human muc
7	22	100.0	1743	14 ADV66231	Adv66231 TRP-like
8	22	100.0	2051	13 ADP25150	Adp25150 PRO polyp
9	22	100.0	2052	4 AAF81753	Aaf81753 Human mem
10	22	100.0	2092	5 AAS72274	Aas72274 DNA encod
11	22	100.0	2092	6 ABL90358	Abl90358 Human pol
12	22	100.0	2094	3 AAA39067	Aaa39067 Human sec
13	22	100.0	2095	6 AAI171699	Aai171699 Human TRP
14	22	100.0	2095	6 ABL40754	Abl40754 Human TLC
15	22	100.0	2095	14 ADV66229	Adv66229 TRP-like
16	22	100.0	2140	12 ADQ24925	Adq24925 Human sof
17	22	100.0	11365	4 AAK73827	Aak73827 Human imm
18	22	100.0	13270	9 ADB84283	Adb84283 Human muc
19	22	100.0	20046	4 AAK73826	Aak73826 Human imm

C 20	20	90.9	349980	6	ABQ81847	Abq81847 Bifidobac
C 21	18.4	83.6	2860	5	AAS65933	Aas65933 DNA encod
C 22	17.4	79.1	682	8	ACA49409	Aca49409 Prokaryot
C 23	17.4	79.1	687	8	ACA51484	Aca51484 Prokaryot
C 24	17.4	79.1	3153	2	AAV19895	Aav19895 Glucodext
C 25	17.2	78.2	963	13	ADT46459	Adt46459 Bacterial
C 26	16.8	76.4	447	13	ADT43994	Adt43994 Bacterial
C 27	16.8	76.4	705	3	AZ50958	Aaz50958 Agrobacte
C 28	16.8	76.4	2505	13	ADU01691	Adu01691 Novel hum
C 29	16.4	74.5	1153	4	ABL16397	Abl16397 Drosophil
C 30	16.4	74.5	1274	5	AAS79165	Aas79165 DNA encod
C 31	16.4	74.5	3208	4	ABL16396	Abl16396 Drosophil
C 32	16.4	74.5	4106	4	ABL17062	Abl17062 Drosophil
C 33	16.4	74.5	110000	11	ADM27081_07	Continuation (8 of
C 34	16.2	73.6	230	14	ADY99718	Ady99718 T_reesei
C 35	16.2	73.6	531	12	ACH70153	Ach70153 Human gen
C 36	16.2	73.6	1035	4	AAD08299	Aad08299 Human sec
C 37	16.2	73.6	1228	11	ADM22430	Adm22430 Rat chole
C 38	16.2	73.6	1339	13	ADT16087	Adt16087 Plant CDN
C 39	16.2	73.6	1638	13	ADT46663	Adt46663 Bacterial
C 40	16.2	73.6	2442	4	ABL08173	Abl08173 Drosophil
C 41	16.2	73.6	2609	13	ADX34421	Adx34421 Plant ful
C 42	16.2	73.6	3825	11	ACL28924	Acl28924 Rice abio
C 43	16.2	73.6	4442	4	ABL08172	Abl08172 Drosophil
C 44	16.2	73.6	4816	10	ADE54538	Ade54538 Rat gene
C 45	16.2	73.6	4816	10	ADE54542	Ade54542 Rat gene

ALIGNMENTS

RESULT 1

AEB28958

ID AEB28958 standard; DNA; 22 BP.

XX AC AEB28958;

XX DT 22-SEP-2005 (first entry)

XX DE Human MCOLN1 gene PCR primer IVS-R SEQ ID NO:2.

XX KW mucolipidosis IV; high throughput screening; mucolipin-1; MCOLN1; PCR;
XX KW primer; ss.

XX OS Homo sapiens.

XX PN US2005153300-A1.

XX PD 14-JUL-2005.

XX PF 09-JAN-2004; 2004US-00754446.

XX PR 09-JAN-2004; 2004US-00754446.

XX PA (QUES-) QUEST DIAGNOSTICS INC.

XX PI Sun W, Hantash F;

XX DR WPI; 2005-521160/53.

XX PT Diagnosing mucolipidosis (ML) IV mutant sequence in nucleic acids, by
amplifying the nucleic acid, detecting amplified product with labeled
oligonucleotide probes via a change in fluorescence which indicates the
presence of an ML IV mutant.

XX Claim 1; SEQ ID NO 2; 15pp; English.

XX PS The invention relates to a method (M1) for detecting the presence of
mucolipidosis IV (ML IV) mutant sequences in a nucleic acid. (M1)
XX CC involves contacting the nucleic acid with oligonucleotide primers and
probes, conducting amplification by temperature cycling and monitoring
the accumulation of amplified nucleic acid by detecting an increase in
donor fluorophore fluorescence or decrease in acceptor fluorophore

CC fluorescence, which indicates the presence of ML IV mutant sequence. Also
 CC described: (1) an oligonucleotide (ON1) comprising a sequence
 CC complementary to the coding or non-coding strand of the mucolipin-1
 CC (MCOLN1) gene; and (2) a kit (K1) for amplifying sequences of ML IV
 CC sequences, comprising ON1. (M1) is useful for determining the presence of
 CC ML IV mutant sequence in a nucleic acid; and/or wild-type sequences. (M1)
 CC is useful for genetic counseling of individuals at risk for ML IV or at
 CC risk to having offspring afflicted with ML IV. The MCOLN1 gene is located
 CC on chromosome 19p13. The present sequence represents a PCR primer for the
 CC MCOLN1 gene, which is used in the exemplification of the present
 CC invention.

XX
 SQ Sequence 22 BP; 7 A; 7 C; 3 G; 5 T; 0 U; 0 Other;
 Query Match 100.0%; Score 22; DB 14; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.3;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TAACCACCATCGGATCAATGTC 22
 |||||
 Db 1 TAACCACCATCGGATCAATGTC 22

RESULT 2
 AAI61022
 ID AAI61022 standard; cDNA; 1619 BP.

XX AAI61022;

DT 22-OCT-2001 (first entry)

XX Human polynucleotide SEQ ID NO 5011.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.

XX 21-JAN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-00552317.

XX 20-JUN-2000; 2000US-00598042.

XX 19-JUL-2000; 2000US-00620312.

XX 03-AUG-2000; 2000US-00653450.

XX 14-SEP-2000; 2000US-00662191.

XX 19-OCT-2000; 2000US-00693036.

XX 29-NOV-2000; 2000US-00727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;

XX Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

XX P-PSDB; AAM41866.

XX Novel nucleic acids and polypeptides, useful for treating disorders such

CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression.
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemia and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification

SQ Sequence 1619 BP; 350 A; 450 C; 518 G; 301 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 4; Length 1619;
 Best Local Similarity 100.0%; Pred. No. 0.51;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TAACCACCATCGGATCAATGTC 22
 |||||
 Db 1382 TAACCACCATCGGATCAATGTC 1403

RESULT 3

AAI71700/c

ID AAI71700 standard; cDNA; 1740 BP.

XX AAI71700;

DT 29-JAN-2002 (first entry)

XX Human TRP-like calcium channel TLCC-2 coding sequence #2.

XX Human; TLCC-2; TRP-like calcium channel; membrane excitability;
 KW nociception; nootropic; neuroprotective; antiparkinsonian; cytostatic;
 KW hypotensive; antidepressant; analgesic; anticonvulsant; tranquiliser;
 KW Parkinson's disease; Huntington's disease; multiple sclerosis;
 KW Gilles de la Tourette's syndrome; autonomic function disorder; cancer;
 KW neuroleptic; gene therapy; Alzheimer's disease; CNS disorder; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..1740

XX /*tag= a

XX /product= "TLCC-2"

XX /partial

XX WO200177331-A1.

XX 18-OCT-2001.

XX 06-APR-2001; 2001WO-US011442.

XX 07-APR-2000; 2000US-00544797.

XX (MILL-) MILLENIUM PHARM INC.

XX Curtis RAJ, Silos-Santiago I;

XX WPI; 2002-010913/01.

XX P-PSDB; AAM51858.

XX Novel isolated human transient receptor potential-like calcium channel

XX protein-2 useful for treating Alzheimer's disease, depression, amnesia,

XX pain disorder, and cancer.

XX Claim 1; Fig 1; 148pp; English.

XX The present invention relates to the protein and coding sequences of

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 16:41:33 ; Search time 531.344 Seconds
(without alignments)
342.389 Million cell updates/sec

Title: US-10-754-446-2

Perfect score: 22

Sequence: 1 TAACCACCATCGGATCAATGTC 22

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	22	9	US-10-754-446-2
C 2	22	100.0	1400	10	US-11-060-756-2850
C 3	22	100.0	1400	10	US-11-060-756-7122
C 4	22	100.0	1740	3	US-09-828-466-3
C 5	22	100.0	1740	5	US-10-103-458-3
C 6	22	100.0	1743	8	US-10-782-695-6
C 7	22	100.0	2051	3	US-09-851-494B-2
C 8	22	100.0	2052	3	US-09-965-529-50
C 9	22	100.0	2052	3	US-09-969-680A-50
C 10	22	100.0	2052	10	US-11-048-692-50
C 11	22	100.0	2092	6	US-10-264-237-920
C 12	22	100.0	2092	9	US-10-450-763-8078
C 13	22	100.0	2094	3	US-09-820-893-26
C 14	22	100.0	2094	7	US-10-607-565-26
C 15	22	100.0	2095	3	US-09-828-466-1
C 16	22	100.0	2095	5	US-10-103-458-1
C 17	22	100.0	2095	8	US-10-782-695-4
C 18	22	100.0	2140	8	US-10-723-860-7745
C 19	22	100.0	13270	3	US-09-851-494B-1
C 20	20	90.9	25	10	US-11-060-756-171643
C 21	20	90.9	2256646	7	US-10-470-565-1
C 22	18.4	83.6	2860	9	US-10-450-763-1737
C 23	17.4	79.1	682	7	US-10-282-122A-37279

```

c 24 17.4 79.1 687 7 US-10-282-122A-39354 Sequence 39354, A
c 25 17.2 78.2 496 7 US-10-424-599-142334 Sequence 142334, A
c 26 17.2 78.2 963 6 US-10-369-493-44897 Sequence 44897, A
c 27 17.2 78.2 2731748 7 US-10-297-465A-1 Sequence 1, Appli
c 28 16.8 76.4 447 6 US-10-369-493-42432 Sequence 42432, A
c 29 16.8 76.4 493 4 US-09-925-065A-364043 Sequence 364043, A
c 30 16.8 76.4 24708 3 US-09-924-101-7 Sequence 7, Appli
c 31 16.4 74.5 25 9 US-10-843-527-25559 Sequence 25559, A
c 32 16.4 74.5 25 9 US-10-843-527-212618 Sequence 212618, A
c 33 16.4 74.5 519 4 US-09-925-065A-133197 Sequence 133197, A
c 34 16.4 74.5 521 4 US-09-925-065A-636380 Sequence 636380, A
c 35 16.4 74.5 1153 10 US-11-097-143-21836 Sequence 21836, A
c 36 16.4 74.5 1274 9 US-10-450-763-14969 Sequence 14969, A
c 37 16.4 74.5 3208 10 US-11-097-143-21835 Sequence 21835, A
c 38 16.4 74.5 4106 10 US-11-097-143-22834 Sequence 22834, A
c 39 16.2 73.6 230 9 US-10-950-009-1072 Sequence 1072, Ap
c 40 16.2 73.6 506 4 US-09-925-065A-235084 Sequence 235084, A
c 41 16.2 73.6 531 6 US-10-029-386-3348 Sequence 3348, Ap
c 42 16.2 73.6 554 4 US-09-925-065A-523439 Sequence 523439, A
c 43 16.2 73.6 554 4 US-09-925-065A-523440 Sequence 523440, A
c 44 16.2 73.6 557 4 US-09-925-065A-940573 Sequence 940573, A
c 45 16.2 73.6 600 9 US-10-972-079-19463 Sequence 19463, A

```

ALIGNMENTS

```

RESULT 1
US-10-754-446-2
; Sequence 2, Application US/10754446
; Publication No. US20050153300A1
; GENERAL INFORMATION:
; APPLICANT: SUN, WEIMIN
; APPLICANT: HANTASH, PERAS
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DETECTION OF
; TITLE OF INVENTION: MUCOLIPIDOSIS IV MUTATIONS
; FILE REFERENCE: 034827-2301
; CURRENT APPLICATION NUMBER: US/10/754, 446
; CURRENT FILING DATE: 2004-01-09
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 3.2
; SEQ ID NO 2
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-10-754-446-2

```

```

Query Match 100.0%; Score 22; DB 9; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 TAACCACCATCGGATCAATGTC 22
|||||
Db 1 TAACCACCATCGGATCAATGTC 22

```

```

RESULT 2
US-11-060-756-2850/C
; Sequence 2850, Application US/11060756
; Publication No. US20050221354A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; TITLE OF INVENTION: Target Genes
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060,756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: Patent In version 3.2
; SEQ ID NO 2850

```

```
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-060-756-2850

Query Match      100.0%; Score 22; DB 10; Length 1400;
Best Local Similarity 100.0%; Pred. No. 0.92;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACCACCATCGGATCAATGTC 22
    |||||
Db 42 TAACCACCATCGGATCAATGTC 21

RESULT 3
US-11-060-756-7122/c
; Sequence 7122, Application US/11060756
; Publication No. US20050221354A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; FILE REFERENCE: Target Genes
; CURRENT APPLICATION NUMBER: US/11/060,756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7122
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-060-756-7122

Query Match      100.0%; Score 22; DB 10; Length 1400;
Best Local Similarity 100.0%; Pred. No. 0.92;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACCACCATCGGATCAATGTC 22
    |||||
Db 42 TAACCACCATCGGATCAATGTC 21

RESULT 4
US-09-828-466-3/c
; Sequence 3, Application US/09828466
; Patent No. US20020035056A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Silos-Santiago, Immaculada
; TITLE OF INVENTION: 54420, A NOVEL HUMAN CALCIUM CHANNEL
; FILE REFERENCE: MNI-125CP
; CURRENT APPLICATION NUMBER: US/09/828,466
; CURRENT FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 09/544,797
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1740)
US-09-828-466-3

Query Match      100.0%; Score 22; DB 3; Length 1740;
Best Local Similarity 100.0%; Pred. No. 0.94;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACCACCATCGGATCAATGTC 22
    |||||
```

```
Db 568 TAACCACCATCGGATCAATGTC 547

RESULT 5
US-10-103-458-3/c
; Sequence 3, Application US/10103458
; Publication No. US20020197680A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; TITLE OF INVENTION: 54420, A NOVEL HUMAN CALCIUM CHANNEL
; FILE REFERENCE: MNI-125
; CURRENT APPLICATION NUMBER: US/10/103,458
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US/09/544,797
; PRIOR FILING DATE: PEIOE FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1740)
US-10-103-458-3

Query Match      100.0%; Score 22; DB 5; Length 1740;
Best Local Similarity 100.0%; Pred. No. 0.94;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACCACCATCGGATCAATGTC 22
    |||||
Db 568 TAACCACCATCGGATCAATGTC 547

RESULT 6
US-10-782-695-6/c
; Sequence 6, Application US/10782695
; Publication No. US20040248160A1
; GENERAL INFORMATION:
; APPLICANT: Gluckemann, Maria A.
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Tsai, Fong-Ying
; APPLICANT: Hodge, Martin R.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: NOVEL 14275, 54420, 8797, 27439, 68730,
; FILE REFERENCE: MPI04-0020NMIM
; CURRENT APPLICATION NUMBER: US/10/782,695
; CURRENT FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: US 09/945,254
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,829
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 09/945,301
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,301
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 10/007,399
; PRIOR FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: US 09/390,039
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: US 09/146,416
; PRIOR FILING DATE: 1998-09-03
; PRIOR APPLICATION NUMBER: US 10/024,036
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/258,222
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 10/103,458
; PRIOR FILING DATE: 2002-03-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 16:59:57 ; Search time 270.688 Seconds
(without alignments)
65.779 Million cell updates/sec

Title: US-10-754-446-2

Perfect score: 22

Sequence: 1 taaccaccatcgatcaatgtc 22

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6038814 seqs, 404674181 residues

Total number of hits satisfying chosen parameters: 12077628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA New:*

1: /cgn2_6/ptodata/2/pubna/US08_NEW_PUB.seq:
2: /cgn2_6/ptodata/2/pubna/US06_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubna/US07_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubna/PCT_NEW_PUB.seq:
5: /cgn2_6/ptodata/2/pubna/US09_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubna/US10_NEW_PUB.seq:
7: /cgn2_6/ptodata/2/pubna/US11_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubna/US11_NEW_PUB.seq:
9: /cgn2_6/ptodata/2/pubna/US11_NEW_PUB.seq:
10: /cgn2_6/ptodata/2/pubna/US60_NEW_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	16.8	76.4	1107	6	US-10-750-185-53745 Sequence 53745, A
C 2	16.8	76.4	1107	6	US-10-750-623-53745 Sequence 53745, A
C 3	16.2	73.6	1458	7	US-11-133-360-4 Sequence 4, Appl
C 4	16.2	73.6	1786	6	US-10-750-185-46643 Sequence 46643, A
C 5	16.2	73.6	1786	6	US-10-750-623-46643 Sequence 46643, A
C 6	15.8	71.8	600	7	US-11-128-061-3932 Sequence 3932, Ap
C 7	15.8	71.8	669	7	US-11-128-061-290 Sequence 290, App
C 8	15.8	71.8	2081	6	US-10-750-185-27573 Sequence 27573, A
C 9	15.8	71.8	2081	6	US-10-750-623-27573 Sequence 27573, A
C 10	15.6	70.9	506	7	US-11-128-061-2800 Sequence 2800, Ap
C 11	15.6	70.9	506	7	US-11-128-061-6442 Sequence 6442, Ap
C 12	15.6	70.9	98345	7	US-11-112-908-36 Sequence 36, Appl
C 13	15.6	70.9	127340	7	US-11-112-908-35 Sequence 35, Appl
C 14	15.6	70.9	157224	7	US-11-112-908-51 Sequence 51, Appl
C 15	15.6	70.9	170189	7	US-11-112-908-50 Sequence 50, Appl
C 16	15.4	70.0	88873	6	US-10-995-561-13383 Sequence 13383, A
C 17	15.4	70.0	146656	7	US-11-121-086-68 Sequence 68, Appl
C 18	15.4	70.0	154452	7	US-11-121-086-74 Sequence 74, Appl
C 19	15.4	70.0	190882	7	US-11-121-086-69 Sequence 69, Appl
C 20	15.2	69.1	25	7	US-11-136-527-311442 Sequence 311442, A
C 21	15.2	69.1	25	7	US-11-136-527-311448 Sequence 311448, A
C 22	15.2	69.1	448	7	US-11-123-896-406 Sequence 406, App
C 23	15.2	69.1	600	7	US-11-136-527-6599 Sequence 6599, App

C 24	15.2	69.1	1870	6	US-10-750-185-38563 Sequence 38563, A
C 25	15.2	69.1	1870	6	US-10-750-623-38563 Sequence 38563, A
C 26	15.2	69.1	3328	7	US-11-136-527-2503 Sequence 2503, Ap
C 27	15	68.2	1400	7	US-11-136-527-4366 Sequence 4366, Ap
C 28	15	68.2	2746	7	US-11-136-527-270 Sequence 270, App
C 29	14.8	67.3	25	7	US-11-136-527-311434 Sequence 311434, A
C 30	14.8	67.3	1509	6	US-10-454-437-211 Sequence 211, App
C 31	14.8	67.3	4091	6	US-10-750-185-32344 Sequence 32344, A
C 32	14.8	67.3	4091	6	US-10-750-623-32344 Sequence 32344, A
C 33	14.8	67.3	182303	7	US-11-121-086-45 Sequence 45, Appl
C 34	14.6	66.4	408	6	US-10-793-626-2675 Sequence 2675, Ap
C 35	14.6	66.4	600	6	US-10-750-185-3617 Sequence 3617, Ap
C 36	14.6	66.4	600	6	US-10-750-623-3617 Sequence 3617, Ap
C 37	14.6	66.4	1161	6	US-10-793-626-2299 Sequence 2299, Ap
C 38	14.6	66.4	1398	7	US-11-133-360-2 Sequence 2, Appl
C 39	14.6	66.4	1401	7	US-11-133-360-6 Sequence 6, Appl
C 40	14.6	66.4	1401	7	US-11-133-360-8 Sequence 8, Appl
C 41	14.6	66.4	1951	6	US-10-750-185-64757 Sequence 64757, A
C 42	14.6	66.4	1951	6	US-10-750-623-64757 Sequence 64757, A
C 43	14.6	66.4	2083	7	US-11-167-856-15 Sequence 15, Appl
C 44	14.6	66.4	2290	6	US-10-750-185-44532 Sequence 44532, A
C 45	14.6	66.4	2290	6	US-10-750-623-44532 Sequence 44532, A

ALIGNMENTS

RESULT 1
US-10-750-185-53745/c
; Sequence 53745, Application US/10750185
; Publication No. US200502603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10750.185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53745
; LENGTH: 1107
; TYPE: DNA
; ORGANISM: Bovine 19866881030294
US-10-750-185-53745

Query Match 76.4%; Score 16.8; DB 6; Length 1107;
Best Local Similarity 90.0%; Pred. No. 45;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TAACCACCATCGATCAATG 20
|||||
DB 800 TAACCACCATCGATCAATG 781
|||||

RESULT 2
US-10-750-623-53745/c
; Sequence 53745, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen

```
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53745
; LENGTH: 1107
; TYPE: DNA
; ORGANISM: Bovine 19866881030294
US-10-750-623-53745

Query Match      76.4%; Score 16.8; DB 6; Length 1107;
Best Local Similarity 90.0%; Pred. No. 45;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TAACCACCATCGGATCAATG 20
Db 800 TAACCACCATCGATCATG 781

RESULT 3
US-11-133-360-4
; Sequence 4, Application US/11133360
; Publication No. US20050287172A1
; GENERAL INFORMATION:
; APPLICANT: Medimmune Vaccines, Inc et al.
; TITLE OF INVENTION: INFLUENZA HEMAGGLUTININ AND NEURAMINIDASE VARIANTS
; FILE REFERENCE: FL260
; CURRENT APPLICATION NUMBER: US/11/133,360
; CURRENT FILING DATE: 2005-05-20
; PRIOR APPLICATION NUMBER: 60/574,553
; PRIOR FILING DATE: 2004-05-25
; PRIOR APPLICATION NUMBER: 60/657,554
; PRIOR FILING DATE: 2005-02-28
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
; LENGTH: 1458
; TYPE: DNA
; ORGANISM: Influenza virus
US-11-133-360-4

Query Match      73.6%; Score 16.2; DB 7; Length 1458;
Best Local Similarity 85.7%; Pred. No. 90;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TAACCACCATCGGATCAATGT 21
Db 40 TAACAACCATGGATCAATCT 60

RESULT 4
US-10-750-185-46643
; Sequence 46643, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
```

```
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46643
; LENGTH: 1786
; TYPE: DNA
; ORGANISM: Bovine 19866881022693
US-10-750-185-46643

Query Match      73.6%; Score 16.2; DB 6; Length 1786;
Best Local Similarity 85.7%; Pred. No. 92;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TAACCACCATCGGATCAATGT 21
Db 882 TGACCACCATCGCATCACTGT 902

RESULT 5
US-10-750-623-46643
; Sequence 46643, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46643
; LENGTH: 1786
; TYPE: DNA
; ORGANISM: Bovine 19866881022693
US-10-750-623-46643

Query Match      73.6%; Score 16.2; DB 6; Length 1786;
Best Local Similarity 85.7%; Pred. No. 92;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TAACCACCATCGGATCAATGT 21
Db 882 TGACCACCATCGCATCACTGT 902

RESULT 6
US-11-128-061-3932
; Sequence 3932, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: MOUNTS, William M.
; APPLICANT: HANN, Louane E.
; APPLICANT: SINACORE, Martin S.
; APPLICANT: LEONARD, Mark W.
; APPLICANT: BROWN, Eugene L.
; APPLICANT: MILLER, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
```


GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 16:32:47 ; Search time 80.08 Seconds
(without alignments)
488.341 Million cell updates/sec

Title: US-10-754-446-2

Perfect score: 22

Sequence: 1 taaccacccatcgatcaatgtc 22

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PCUTUS COMB.seq.*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	22	100.0	2051	3	US-09-949-016-1712 Sequence 1712, Ap
C 2	22	100.0	15353	3	US-09-949-016-13454 Sequence 13454, A
C 3	16	72.7	1650	3	US-09-107-532A-3186 Sequence 3186, Ap
C 4	15.8	71.8	537	3	US-09-489-039A-5168 Sequence 5168, Ap
C 5	15.8	71.8	572	3	US-09-533-559-6112 Sequence 6112, Ap
C 6	15.8	71.8	601	3	US-09-949-016-79854 Sequence 79854, A
C 7	15.8	71.8	601	3	US-09-949-016-83923 Sequence 83923, A
C 8	15.8	71.8	1614	2	US-08-867-941-10 Sequence 10, Appl
C 9	15.8	71.8	1614	2	US-08-867-941-10 Sequence 10, Appl
C 10	15.8	71.8	1623	2	US-08-867-941-5 Sequence 5, Appli
C 11	15.8	71.8	1623	2	US-08-867-941-5 Sequence 5, Appli
C 12	15.8	71.8	1635	3	US-09-540-236-840 Sequence 840, App
C 13	15.8	71.8	4657	3	US-09-254-325-1 Sequence 1, Appli
C 14	15.8	71.8	5643	3	US-09-079-415-5 Sequence 5, Appli
C 15	15.8	71.8	5643	3	US-08-750-458A-1 Sequence 1, Appli
C 16	15.8	71.8	7641	2	US-08-867-941-6 Sequence 6, Appli
C 17	15.8	71.8	7641	3	US-09-074-658-6 Sequence 6, Appli
C 18	15.8	71.8	7650	2	US-08-867-941-1 Sequence 1, Appli
C 19	15.8	71.8	7650	3	US-09-074-658-1 Sequence 1, Appli
C 20	15.8	71.8	41322	3	US-10-024-396-13 Sequence 13, Appl
C 21	15.8	71.8	49617	3	US-09-596-002-28 Sequence 28, Appl
C 22	15.8	71.8	90472	3	US-09-949-016-14038 Sequence 14038, A
C 23	15.8	71.8	96340	3	US-09-949-016-15863 Sequence 15863, A
C 24	15.8	71.8	455726	3	US-09-949-016-14157 Sequence 14157, A

C 25	15.8	71.8	481115	3	US-09-949-016-11940 Sequence 11940, A
C 26	15.6	70.9	762	3	US-09-252-991A-3481 Sequence 3481, Ap
C 27	15.6	70.9	816	3	US-09-489-039A-4176 Sequence 4176, Ap
C 28	15.6	70.9	834	3	US-09-489-039A-2174 Sequence 2174, Ap
C 29	15.6	70.9	1203	3	US-09-252-991A-3468 Sequence 3468, Ap
C 30	15.6	70.9	1269	3	US-09-252-991A-3470 Sequence 3470, Ap
C 31	15.6	70.9	1428	3	US-09-248-796A-3428 Sequence 3428, Ap
C 32	15.6	70.9	1697	3	US-09-270-767-10255 Sequence 10255, A
C 33	15.6	70.9	1876	3	US-08-192-271-8 Sequence 8, Appli
C 34	15.6	70.9	1876	3	PCT-US94-01572-8 Sequence 8, Appli
C 35	15.6	70.9	2231	3	US-09-949-016-4932 Sequence 4932, Ap
C 36	15.6	70.9	2231	3	US-09-949-016-4933 Sequence 4933, Ap
C 37	15.6	70.9	3041	3	US-10-126-120-1 Sequence 1, Appli
C 38	15.6	70.9	4081	3	US-08-939-774A-1 Sequence 1, Appli
C 39	15.6	70.9	4558	2	US-08-309-512-2 Sequence 2, Appli
C 40	15.6	70.9	4558	6	PCT-US92-08756A-2 Sequence 2, Appli
C 41	15.6	70.9	7400	2	US-07-674-852-1 Sequence 1, Appli
C 42	15.6	70.9	7400	3	US-08-473-185-1 Sequence 1, Appli
C 43	15.6	70.9	7400	3	US-09-171-387-3 Sequence 3, Appli
C 44	15.6	70.9	113379	3	US-09-949-016-17561 Sequence 17561, A
C 45	15.6	70.9	113379	3	US-09-949-016-17562 Sequence 17562, A

ALIGNMENTS

RESULT 1

US-09-949-016-1712/c
; Sequence 1712, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1712
; LENGTH: 2051
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1712

Query Match 100.0%; Score 22; DB 3; Length 2051;

Best Local Similarity 100.0%; Pred. No. 0.24; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACACCATCGGATCAATGTC 22

DB 693 TAACACCATCGGATCAATGTC 672

RESULT 2

US-09-949-016-13454/c
; Sequence 13454, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13454
; LENGTH: 15353
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13454

Query Match 100.0%; Score 22; DB 3; Length 15353;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACCACCATCGGATCAATGTC 22
Db 6298 TAACCACCATCGGATCAATGTC 6277

RESULT 3

US-09-107-532A-3186/c
; Sequence 3186, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Walcham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354

; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 3186:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1650 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8) LOCATION 1...1650
; SEQUENCE DESCRIPTION: SEQ ID NO: 3186:

US-09-107-532A-3186

Query Match 72.7%; Score 16; DB 3; Length 1650;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AACCCACCATCGGATCA 17
Db 578 AACCCACCATCGGATCA 563

RESULT 4

US-09-489-039A-5168/c
; Sequence 5168, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 5168
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5168

Query Match 71.8%; Score 15.8; DB 3; Length 537;
Best Local Similarity 89.5%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TAACCACCATCGGATCAAT 19
Db 151 TAACCACCATCGATTCAT 133

RESULT 5

US-09-533-559-6112
; Sequence 6112, Application US/09533559
; Patent No. 6902887
; GENERAL INFORMATION:

; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groth Clausen
; APPLICANT: Peter Bjarke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; EXPRESSION

; FILE REFERENCE: 5849.200-US
; CURRENT APPLICATION NUMBER: US/09/533,559
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 09/273,623
; EARLIER FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6112
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Aspergillus oryzae

; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(572)
; OTHER INFORMATION: n = A,T,C or G
US-09-533-559-6112

Query Match 71.8%; Score 15.8; DB 3; Length 572;
Best Local Similarity 89.5%; Pred. No. 2.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 15:49:05 ; Search time 736.68 Seconds
(without alignments)
1157.425 Million cell updates/sec

Title: US-10-754-446-1

Perfect score: 15

Sequence: 1 agcggcgagactca 15

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_ats.*
11: gb_av.*
12: gb_un.*
13: gb_vi.*
14: gb_htg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15	100.0	790	AF30557283	AF305574 Homo sapi
2	15	100.0	13270	AF287270	AF287270 Homo sapi
3	15	100.0	21081	AF106590	AF106590 Caenorhab
4	15	100.0	36859	AF003141	AF003141 Caenorhab
5	15	100.0	103303	AY3339213	AY3339213 Takifugu
6	15	100.0	110000	1	Continuation (4 of
7	15	100.0	155645	AC021153	AC021153 Homo sapi
8	15	100.0	166981	AC157751	AC157751 Callithri
9	15	100.0	173126	AC008878	AC008878 Homo sapi
10	15	100.0	180213	AC150812	AC150812 Callithri
11	15	100.0	197202	AC153077	AC153077 Callithri
12	15	100.0	208232	9	AC149085 Mus muscu
13	15	100.0	221553	9	AC149282 Mus muscu
14	15	100.0	228698	14	AC127479 Mus muscu
15	15	100.0	237653	14	AC073760 Mus muscu
16	15	100.0	291000	1	SC0939105
17	14	93.3	563	15	AJ836913 Arabidops
18	14	93.3	789	9	AF296282 Mus muscu

19	14	93.3	789	9	AF296283	AF296283 Mus muscu
c 20	14	93.3	1250	9	S70756S1	S70756 Tlx-1-homoe
c 21	14	93.3	1950	6	AR622581	AR622581 Sequence
22	14	93.3	2262	9	AB127583	AB127583 Mus muscu
23	14	93.3	2428	8	AK096823	AK096823 Homo sapi
c 24	14	93.3	2685	9	BC059941	BC059941 Mus muscu
c 25	14	93.3	3238	5	AF427524	AF427524 Danio rer
c 26	14	93.3	4102	6	CQ986467	CQ986467 Sequence
c 27	14	93.3	10246	9	MMU251787	AJ251787 Mus muscu
c 28	14	93.3	10586	1	AE004991	AE004991 Halobacte
c 29	14	93.3	10696	1	AE012476	AE012476 Xanthomon
c 30	14	93.3	20074	8	HSA9794	AJ009794 Homo sapi
c 31	14	93.3	20232	6	AR619802	AR619802 Sequence
c 32	14	93.3	45653	5	EX323016	BX323016 Zebrafish
c 33	14	93.3	51099	8	AL357395	AL357395 Human DNA
c 34	14	93.3	53425	14	AY211534	AY211534 Zea mays
c 35	14	93.3	89015	14	AC008458	AC008458 Homo sapi
c 36	14	93.3	94092	14	AC155398	AC155398 Zea mays
c 37	14	93.3	110000	1	AE016822_15	Continuation (16 o
c 38	14	93.3	110000	1	AE017282_18	Continuation (19 o
c 39	14	93.3	110000	1	AE017282_29	Continuation (30 o
c 40	14	93.3	110000	1	AE017340_16	Continuation (17 o
41	14	93.3	110000	1	AP006618_44	Continuation (45 o
42	14	93.3	110000	1	AP006618_45	Continuation (46 o
c 43	14	93.3	110000	1	AP006618_58	Continuation (59 o
c 44	14	93.3	110000	1	AY596297_08	Continuation (9 of
45	14	93.3	110000	1	AY596297_28	Continuation (29 o

ALIGNMENTS

RESULT 1
AF30557283
LOCUS Homo sapiens mucopolin 1 (MCOLN1) gene, exons 3 and 4.
DEFINITION AF305574
ACCESSION AF305574.1 GI:11991200
VERSION 1 (bases 1 to 790)
KEYWORDS 3 of 8
SEGMENT Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 790)
AUTHORS Bargal, R., Avidan, N., Ben-Asher, E., Olender, Z., Zeigler, M., Frumkin, A., Raas-Rothschild, A., Glusman, G., Lancet, D. and Bach, G.
TITLE Identification of the gene causing mucopolidiosis type IV
JOURNAL Nat. Genet. 26 (1), 118-123 (2000)
PUBMED 10973263
REFERENCE 2 (bases 1 to 790)
AUTHORS Bargal, R., Avidan, N., Ben-Asher, E., Olender, A., Zeigler, M., Frumkin, A., Raas-Rothschild, A., Glusman, G., Lancet, D. and Bach, G.
TITLE Direct Submission
JOURNAL Submitted (13-SEP-2000) Molecular Genetics, The Weizmann Institute of Science, P. O. Box 26, Rehovot 76100, Israel
FEATURES
source 1. .790
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
267. .434
/gene="MCOLN1"
/number=3
589. .757
/gene="MCOLN1"
/number=4

ORIGIN
Query Match 100.0%; Score 15; DB 8; Length 790;
Best Local Similarity 100.0%; Pred. No. 7.3e+03; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0;

```

Qy      1 AGCGGCGCGACTCA 15
Db      548 AGCGGCGCGACTCA 562

RESULT 2
AF287270      13270 bp      DNA      linear      PRI 30-OCT-2000
LOCUS      Homo sapiens mucolipin (MCOLN1) gene, complete cds.
DEFINITION
ACCESSION      AF287270
VERSION      AF287270.1 GI:9844925
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1 (bases 1 to 13270)
Sun, M., Goldin, E., Stahl, S., Falardeau, J.L., Kennedy, J.C.,
Acierno, J.S. Jr., Bove, C., Kanecki, C.R., Nagle, J., Bromley, M.C.,
Colman, M., Schiffmann, R. and Slaugenaupt, S.A.
Mucopolidosis type IV is caused by mutations in a gene encoding a
novel transient receptor potential channel
Hum. Mol. Genet. 9 (17), 2471-2478 (2000)
11030752
2 (bases 1 to 13270)
Slaugenaupt, S.A.
Direct Submission
JOURNAL
PUBMED
AUTHORS
TITLE
JOURNAL
REFERENCE
1. 13270
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19p13.3"
1401..12753
/gene="MCOLN1"
join(1401..1556,3736..3941,5214..5381,5536..5701,
6295..6403,6639..6735,6933..7032,7372..7478,7596..7745,
7876..7977,8365..8487,9061..9276,12298..12428,
12534..12753)
/gene="MCOLN1"
/product="mucolipin"
join(1526..1556,3736..3941,5214..5381,5536..5701,
6295..6403,6639..6735,6933..7032,7372..7478,7596..7745,
7876..7977,8365..8487,9061..9276,12298..12428,
12534..12570)
/gene="MCOLN1"
/codon_start=1
/product="mucolipin"
/protein_id="AAG00798.1"
/db_xref="GI:9844926"
/translators="MTAPAGRGSETERLLTPNGYQAGSPAPPTPEEDLRER
LKYPFSPCKPRAKGRKPKMLQVILVTVQLLGLSNQLAVFRENTIAPR
HLFLYSGDADTDFAITREQLQALFHAVDQLPDLVSLGAYVRGGDPWTNG
SGALCORYYHGHVDPAITDPMVVDICIQDPPPPPPPSDGLTLLESY
KNLTLPKLVNTIHPRLTINLQSLINNEIDCYTFVSLITFDKHAHSGRIPISLE
TOAHIOCKHPSPFHQDNGSFRLLFDVVVILTCSLFLCARSLRGLLQNFVGFWM
WRQGRVISLWELWGVYLLVTSVLTISGTMKIGIEAKNLASVDVCSILLGT
STLLVWGVIRYLTFPHNYLLIATLRVLPSPVNRFCCEVAVIYLVGFCGNIPLGY
HVFRLSUSWSECLFSLNGDMFVTPFAAQGRSSLVWFLSGLYLSFISLFYM
VLSLFTALITGYDTIKHPGAGAESELAQYTAQQDSPTSGKFRRSGSACSLCC
CGRDPSEHSLVN"

Qy      1 AGCGGCGCGACTCA 15
Db      5495 AGCGGCGCGACTCA 5509

RESULT 3
AF106590/c      21081 bp      DNA      linear      INV 21-SEP-2004
LOCUS      Caenorhabditis elegans cosmid F08B6, complete sequence.
DEFINITION
ACCESSION      AF106590
VERSION      AF106590.1 GI:3886091
KEYWORDS      HTG.
SOURCE
ORGANISM      Caenorhabditis elegans
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditidae; Rhabditidae; Feloderinae; Caenorhabditis.
1 (bases 1 to 21081)
WormBase Consortium
Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium
Science 282 (5396), 2012-2018 (1998)
9851916
2 (bases 1 to 21081)
Le, T.T., Scheet, P. and Kemp, K.
The sequence of C. elegans cosmid F08B6
Unpublished (2001)
3 (bases 1 to 21081)
Waterston, R.
Direct Submission
Submitted (15-NOV-1998) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
4 (bases 1 to 21081)
Waterston, R.
Direct Submission
Submitted (28-JUN-2001) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
5 (bases 1 to 21081)
Waterston, R.
Direct Submission
Submitted (22-AUG-2001) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
6 (bases 1 to 21081)
Waterston, R.
Direct Submission
Submitted (14-SEP-2001) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
7 (bases 1 to 21081)
Waterston, R.
Direct Submission
Submitted (05-OCT-2001) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
8 (bases 1 to 21081)
Waterston, R.
Direct Submission
Submitted (22-MAY-2002) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
9 (bases 1 to 21081)
Waterston, R.
Direct Submission
Submitted (19-NOV-2002) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
10 (bases 1 to 21081)
Wilson, R.
Direct Submission
Submitted (15-JUN-2003) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.

```

ORIGIN

Query Match 100.0%; Score 15; DB 8; Length 13270;
 Best Local Similarity 100.0%; Pred. No. 4e+03;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 15:55:50 ; Search time 1677.48 Seconds
(without alignments)
418.369 Million cell updates/sec

Title: US-10-754-446-1

Perfect score: 15

Sequence: 1 agcgggcccggactca 15

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1.*

2: gb_est2.*

3: gb_est3.*

4: gb_hic.*

5: gb_est4.*

6: gb_est5.*

7: gb_est6.*

8: gb_est7.*

9: gb_ges1.*

10: gb_ges2.*

11: gb_ges3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	15	100.0	301	5	BY345580 BY345580
C 2	15	100.0	331	5	BY345176 BY345176
C 3	15	100.0	338	5	BY102360 BY102360
C 4	15	100.0	375	5	BY087147 BY087147
C 5	15	100.0	375	5	C08554 C08554 YUji
C 6	15	100.0	378	5	BY299457 BY299457
C 7	15	100.0	379	5	C71190 C71190 YUji
C 8	15	100.0	387	9	CC423137 PUEHT72TD
C 9	15	100.0	469	1	AV631303 AV631303
C 10	15	100.0	480	5	BY249487 BY249487
C 11	15	100.0	502	7	CK583489 CK583489
C 12	15	100.0	519	7	CK583423 CK583423
C 13	15	100.0	568	7	CK577649 CK577649
C 14	15	100.0	624	3	BY726669 BY726669
C 15	15	100.0	638	1	AU200865 AU200865
C 16	15	100.0	779	10	CG116594 CG116594
C 17	15	100.0	786	10	AG602966 AG602966
C 18	15	100.0	797	7	CN200405 CN200405
C 19	15	100.0	832	8	CK892033 CK892033
C 20	15	100.0	962	6	CA464646 CA464646
C 21	15	100.0	1040	3	BQ278046 BQ278046
C 22	15	93.3	194	3	BQ082114 BQ082114

23	14	93.3	208	2	BB568361
24	14	93.3	287	9	AZ121792
C 25	14	93.3	301	1	A1029662
26	14	93.3	306	2	BB419855
27	14	93.3	350	8	W66660
C 28	14	93.3	354	5	BY312288
29	14	93.3	356	8	W99177
30	14	93.3	360	5	C45921
31	14	93.3	374	5	C43796
C 32	14	93.3	389	8	DN368435
33	14	93.3	402	11	DR18114S
C 34	14	93.3	409	5	BY237451
35	14	93.3	436	9	AZ251529
36	14	93.3	455	1	AA050210
37	14	93.3	462	5	BX527893
C 38	14	93.3	469	7	CJ034986
39	14	93.3	471	8	W99993
C 40	14	93.3	473	8	DN367700
C 41	14	93.3	484	5	BX635951
C 42	14	93.3	487	10	CG732852
43	14	93.3	489	8	DN481676
44	14	93.3	489	10	CG465566
C 45	14	93.3	491	2	BE943858

ALIGNMENTS

RESULT 1
BY345580/c 301 bp mRNA linear EST 12-DEC-2002
LOCUS
DEFINITION
BY345580 RIKEN full-length enriched, whole joints Mus musculus cDNA
clone L730015A09 5', mRNA sequence.
ACCESSION
BY345580
VERSION
BY345580.1 GI:26575068
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 301)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nakado, Y., Osato, N., Saito, R., Suzuki, H., Yamada, H.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojibori, T., Baldarelli, R., Hill, D. P., Bult, C.,
Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V.,
Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A.,
Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T.,
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A.,
Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R.,
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J. C., Reed, J., Reid, J., Ring, B. Z., Ringwald, M.,
Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tonita, M.,
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Landier, E. S.,
Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
JOURNAL
PUBMED
12466851

COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,
Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,
Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,
Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and
Hayashizaki, Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Vassilis Aidinis (Biomedical Sciences
Research Center 'Al. Fleming' Institute of Immunology 14-16 Al.
Fleming street 16672 Vari, Greece) whose assistance we gratefully
acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.

FEATURES

source

```
Location/Qualifiers
1. .331
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="L730015A09"
/tissue_type="whole joints"
/clone_lib="RIKEN full-length enriched, whole joints"
```

ORIGIN

```
Query Match      100.0%; Score 15; DB 5; Length 301;
Best Local Similarity 100.0%; Pred. No. 2e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AGCGGGCGGACTCA 15
```

```
Db 17 AGCGGGCGGACTCA 3
```

RESULT 2

```
BY345176/c
```

```
LOCUS
```

```
DEFINITION
```

```
BY345176 RIKEN full-length enriched, whole joints Mus musculus cDNA
```

```
clone L730012A07 5', mRNA sequence.
```

```
ACCESSION
```

```
VERSION
```

```
KEYWORDS
```

```
SOURCE
```

```
ORGANISM
```

```
Mus musculus
```

```
Mus musculus (house mouse)
```

```
REFERENCE
```

```
AUTHORS
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Muridae; Murinae; Mus.
1 (bases 1 to 331)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikado, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,

FEATURES

source

```
Location/Qualifiers
1. .331
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="L730012A07"
/tissue_type="whole joints"
/clone_lib="RIKEN full-length enriched, whole joints"
```

ORIGIN

Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Bruscia, V.,
Catholia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmel, S.,
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawai, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
Maltas, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, J.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,
Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,
Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,
Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and
Hayashizaki, Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Tissues were provided by Vassilis Aidinis (Biomedical Sciences
Research Center 'Al. Fleming' Institute of Immunology 14-16 Al.
Fleming street 16672 Vari, Greece) whose assistance we gratefully
acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 14:15:14 ; Search time 203.52 Seconds
(without alignments)
491.207 Million cell updates/sec

Title: US-10-754-446-1

Perfect score: 15
Sequence: 1 agcggcgccgactca 15

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001as.*
5: Geneseq2001bs.*
6: Geneseq2002as.*
7: Geneseq2002bs.*
8: Geneseq2003as.*
9: Geneseq2003bs.*
10: Geneseq2003cs.*
11: Geneseq2003ds.*
12: Geneseq2004as.*
13: Geneseq2004bs.*
14: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15	100.0	15	14	AEB28957 Human MCO
2	15	100.0	11365	4	AAK73827 Human imm
3	15	100.0	13270	9	ADB84283 Human muc
4	15	100.0	20046	4	AAK73826 Human imm
5	14	93.3	318	6	ABQ91193 M. capsul
6	14	93.3	432	12	ADL84655 DNA up-re
7	14	93.3	432	12	ADL84656 DNA up-re
8	14	93.3	837	13	ADSG1839 Bacterial
9	14	93.3	1950	14	ACL67519 M. xanthu
10	14	93.3	2378	4	AAH23696 Human tum
11	14	93.3	4102	14	ADM10212 Colon pro
12	14	93.3	20232	14	ACL64734 M. xanthu
13	14	93.3	28906	4	AAH23705 Human tum
14	13.4	89.3	97	3	AAH17228 Human sec
15	13.4	89.3	163	2	AAQ75814 AAV ITR.
16	13.4	89.3	252	12	ACH93047 Human gen
17	13.4	89.3	303	6	ABK78538 Bacillus
18	13.4	89.3	365	3	AAA43414 Rat secre
19	13.4	89.3	399	5	AAF66900 Novel hum

20	13.4	89.3	429	4	AAK60192	Human imm
21	13.4	89.3	471	6	ABQ90279	M. capsul
22	13.4	89.3	541	12	ADP81041	Human ova
23	13.4	89.3	556	12	ACH79347	Human gen
c 24	13.4	89.3	616	3	AAF08076	Fusarium
c 25	13.4	89.3	616	13	ADU52117	Fusarium
c 26	13.4	89.3	616	14	ADZ90120	Fusarium
27	13.4	89.3	648	13	ADSG3824	Bacterial
28	13.4	89.3	648	13	ADSG4203	Bacterial
29	13.4	89.3	651	13	ADSG3455	Bacterial
30	13.4	89.3	681	9	ACD10983	Human col
31	13.4	89.3	832	12	ADP81042	Human ova
32	13.4	89.3	930	5	AAH23813	Human tra
33	13.4	89.3	940	4	AAH23813	Human tra
34	13.4	89.3	968	3	AAH23813	Human tra
35	13.4	89.3	968	13	ADU59051	Trichoder
36	13.4	89.3	968	14	ADZ97054	Trichoder
c 37	13.4	89.3	981	11	ACH94304	Klebsiell
c 38	13.4	89.3	1052	5	AAH84354	DNA encod
c 39	13.4	89.3	1152	13	ADT48262	Bacterial
40	13.4	89.3	1281	11	ABD00899	Klebsiell
41	13.4	89.3	1500	4	ABL25207	Drosophil
42	13.4	89.3	1626	13	ADSG59288	Bacterial
43	13.4	89.3	1676	14	ADY17906	DNA encod
44	13.4	89.3	1785	12	ADQ07773	Mouse pol
45	13.4	89.3	1809	12	ADQ07771	Human pol

ALIGNMENTS

RESULT 1

AEB28957
ID AEB28957 standard; DNA; 15 BP.

XX AEB28957;

DT 22-SEP-2005 (first entry)

XX Human MCOLN1 gene PCR primer IVS-F SEQ ID NO:1.

KW mucolipidosis IV; high throughput screening; mucolipin-1; MCOLN1; PCR;
KW primer; ss.

XX Homo sapiens.

OS US2005153300-A1.

PN 14-JUL-2005.

PD 09-JAN-2004; 2004US-00754446.

PF 09-JAN-2004; 2004US-00754446.

PR (QUES-) QUEST DIAGNOSTICS INC.

XX Sun W, Hantash F;

XX WPI; 2005-521160/53.

XX Diagnosing mucolipidosis (ML) IV mutant sequence in nucleic acids, by
PT amplifying the nucleic acid, detecting amplified product with labeled
PT oligonucleotide probes via a change in fluorescence which indicates the
PT presence of an ML IV mutant.

XX Claim 1; SEQ ID NO 1; 15pp; English.

XX The invention relates to a method (M1) for detecting the presence of
XX mucolipidosis IV (ML IV) mutant sequences in a nucleic acid. (M1)
CC involves contacting the nucleic acid with oligonucleotide primers and
CC probes, conducting amplification by temperature cycling and monitoring
CC the accumulation of amplified nucleic acid by detecting an increase in
CC donor fluorophore fluorescence or decrease in acceptor fluorophore

CC fluorescence, which indicates the presence of ML IV mutant sequence. Also
CC described: (1) an oligonucleotide (ON1) comprising a sequence
CC complementary to the coding or non-coding strand of the mucolin-1
CC (MCOLN1) gene; and (2) a kit (K1) for amplifying sequences of ML IV
CC sequences, comprising ON1. (M1) is useful for determining the presence of
CC ML IV mutant sequence in a nucleic acid, and/or wild-type sequences. (M1)
CC is useful for genetic counseling of individuals at risk for ML IV or at
CC risk to having offspring afflicted with ML IV. The MCOLN1 gene is located
CC on chromosome 19p13. The present sequence represents a PCR primer for the
CC MCOLN1 gene, which is used in the exemplification of the present
CC invention.

XX
SQ Sequence 15 BP; 3 A; 5 C; 6 G; 1 T; 0 U; 0 Other;

Query Match 100.0%; Score 15; DB 14; Length 15;

Best Local Similarity 100.0%; Pred. No. 4.1e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCGGGCCGGACTCA 15

Db 1 AGCGGGCCGGACTCA 15

RESULT 2

AAK73827

ID AAK73827 standard; DNA; 11365 BP.

XX

AC AAK73827;

XX

DT 07-NOV-2001 (first entry)

XX

DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:28639.

XX

KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

KW

Cytostatic; Gene therapy; vaccine; metastasis; db.

XX

OS Homo sapiens.

XX

PN WO200157182-A2.

XX

PD 09-AUG-2001.

XX

PF 17-JAN-2001; 2001WO-US0001354.

XX

PR 31-JAN-2000; 2000US-0179065P.

PR

PR 04-FEB-2000; 2000US-0180628P.

PR

PR 24-FEB-2000; 2000US-0184664P.

PR

PR 02-MAR-2000; 2000US-0186350P.

PR

PR 16-MAR-2000; 2000US-0189874P.

PR

PR 17-MAR-2000; 2000US-0190076P.

PR

PR 18-APR-2000; 2000US-0198123P.

PR

PR 19-MAY-2000; 2000US-0205515P.

PR

PR 07-JUN-2000; 2000US-0209467P.

PR

PR 28-JUN-2000; 2000US-0214886P.

PR

PR 30-JUN-2000; 2000US-0215135P.

PR

PR 07-JUL-2000; 2000US-0216647P.

PR

PR 07-JUL-2000; 2000US-0216880P.

PR

PR 11-JUL-2000; 2000US-0217487P.

PR

PR 11-JUL-2000; 2000US-0217496P.

PR

PR 14-JUL-2000; 2000US-0218290P.

PR

PR 26-JUL-2000; 2000US-0220963P.

PR

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226799P.
PR 22-AUG-2000; 2000US-0226811P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239355P.
PR 13-OCT-2000; 2000US-0239377P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 16:41:33 ; Search time 362.28 Seconds
(without alignments)
342.389 Million cell updates/sec

Title: US-10-754-446-1

Perfect score: 15

Sequence: 1 AGCGGCGCGACTCA 15

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA Main.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15	100.0	15	9	US-10-754-446-1 ✓
2	15	100.0	356	8	US-10-425-115-103385
3	15	100.0	13270	3	US-09-851-494B-1
4	14	93.3	201	8	US-10-719-993-29684
5	14	93.3	432	7	US-10-430-201-1048
6	14	93.3	432	7	US-10-430-201-1049
7	14	93.3	527	5	US-10-027-632-130943
8	14	93.3	527	5	US-10-027-632-130944
9	14	93.3	527	6	US-10-027-632-130943
10	14	93.3	527	6	US-10-027-632-130944
11	14	93.3	606	7	US-10-767-701-524
12	14	93.3	837	6	US-10-369-493-37513
13	14	93.3	858	8	US-10-425-115-162360
14	14	93.3	4102	8	US-10-602-494-17
15	14	93.3	19265	8	US-10-719-993-6839
16	13.4	89.3	63	7	US-10-362-906-3
17	13.4	89.3	252	6	US-10-029-386-26242
18	13.4	89.3	303	3	US-09-974-300-5829
19	13.4	89.3	329	8	US-10-425-115-122740
20	13.4	89.3	399	9	US-10-779-543-12575
21	13.4	89.3	438	7	US-10-437-963-5136
22	13.4	89.3	438	8	US-10-425-115-29829
23	13.4	89.3	483	4	US-09-925-065A-755757

C 24	13.4	89.3	499	4	US-09-925-065A-762110	Sequence 762110, A
25	13.4	89.3	556	6	US-10-029-386-12542	Sequence 12542, A
C 26	13.4	89.3	610	5	US-10-027-632-99325	Sequence 99325, A
C 27	13.4	89.3	610	8	US-10-027-632-99325	Sequence 99325, A
C 28	13.4	89.3	616	8	US-10-653-047-599	Sequence 599, App
29	13.4	89.3	638	8	US-10-425-115-140595	Sequence 140595, A
30	13.4	89.3	645	5	US-10-027-632-28813	Sequence 28813, A
31	13.4	89.3	645	5	US-10-027-632-28814	Sequence 28814, A
32	13.4	89.3	645	6	US-10-027-632-28813	Sequence 28813, A
33	13.4	89.3	645	6	US-10-027-632-28814	Sequence 28814, A
34	13.4	89.3	648	6	US-10-369-493-39498	Sequence 39498, A
35	13.4	89.3	648	6	US-10-369-493-39877	Sequence 39877, A
36	13.4	89.3	651	6	US-10-369-493-39129	Sequence 39129, A
37	13.4	89.3	861	5	US-10-001-883-38	Sequence 38, Appl
38	13.4	89.3	861	6	US-10-156-761-2634	Sequence 2634, Ap
39	13.4	89.3	968	8	US-10-653-047-7533	Sequence 7533, Ap
40	13.4	89.3	984	6	US-10-156-761-1305	Sequence 1305, Ap
C 41	13.4	89.3	1052	9	US-10-450-763-20158	Sequence 20158, A
C 42	13.4	89.3	1152	6	US-10-369-493-46700	Sequence 46700, A
43	13.4	89.3	1353	6	US-10-156-761-4728	Sequence 4728, Ap
44	13.4	89.3	1450	8	US-10-425-115-140634	Sequence 140634, A
45	13.4	89.3	1500	10	US-11-097-143-35051	Sequence 35051, A

ALIGNMENTS

RESULT 1
US-10-754-446-1
; Sequence 1, Application US/10754446
; Publication No. US20050153300A1
; GENERAL INFORMATION:
; APPLICANT: HANTASH, FERAS
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DETECTION OF
; FILE REFERENCE: 034827-2301
; CURRENT APPLICATION NUMBER: US/10/754,446
; CURRENT FILING DATE: 2004-01-09
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 1
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-10-754-446-1

Query Match 100.0%; Score 15; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGCGGCGCGACTCA 15
| | | | | | | | | | | | | | | | | | | | | |
Db 1 AGCGGCGCGACTCA 15

RESULT 2
US-10-425-115-103385
; Sequence 103385, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326

```
; SEQ ID NO 103385
; LENGTH: 356
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(356)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_25792C.1
US-10-425-115-103385

Query Match      100.0%; Score 15; DB 8; Length 356;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AGCGGGCGGACTCA 15
Db      255 AGCGGGCGGACTCA 269
|||||

RESULT 3
US-09-851-494B-1
; Sequence 1, Application US/09851494B
; Publication No. US20030064363A1
; GENERAL INFORMATION:
; APPLICANT: ML4 Foundation
; APPLICANT: Goldin, Ehud
; APPLICANT: Staugenhaupt, Susan A.
; APPLICANT: Sun, Mei
; APPLICANT: Agierno, James S.
; TITLE OF INVENTION: A Gene Encoding A New TRP Channel is Mutated in Mucopolipidosis IV
; FILE REFERENCE: 3394/1H557U51
; CURRENT APPLICATION NUMBER: US/09/851.494B
; CURRENT FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 13270
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-851-494B-1

Query Match      100.0%; Score 15; DB 3; Length 13270;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AGCGGGCGGACTCA 15
Db      5495 AGCGGGCGGACTCA 5509
|||||

RESULT 4
US-10-719-993-29684/c
; Sequence 29684, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719.993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29684
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-719-993-29684

Query Match      93.3%; Score 14; DB 8; Length 201;
Best Local Similarity 100.0%; Pred. No. 7.1e+02;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 GCGGGCGGACTCA 15
Db      159 GCGGGCGGACTCA 146
|||||

RESULT 5
US-10-430-201-1048
; Sequence 1048, Application US/10430201
; Publication No. US20040162679A1
; GENERAL INFORMATION:
; APPLICANT: Li, Linheng
; TITLE OF INVENTION: Method for Predicting Gene Potential and Cell Commitment
; FILE REFERENCE: 40716 (1p-010)
; CURRENT APPLICATION NUMBER: US/10/430.201
; CURRENT FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: US 60/370,114
; PRIOR FILING DATE: 2002-05-03
; NUMBER OF SEQ ID NOS: 4879
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1048
; LENGTH: 432
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (212)..(212)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (215)..(215)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (219)..(219)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (232)..(233)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (237)..(237)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (243)..(243)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (254)..(254)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (259)..(259)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (263)..(263)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (309)..(309)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (326)..(326)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (355)..(355)
; OTHER INFORMATION: n is a, c, g, or t
```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 16:59:57 ; Search time 184.56 Seconds
(without alignments)
65.779 Million cell updates/sec

Title: US-10-754-446-1
Perfect score: 15
Sequence: 1 agcggccggactca 15

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 6038814 seqs, 404674181 residues

Total number of hits satisfying chosen parameters: 12077628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA New:*
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	13.4	89.3	20	6	US-10-310-914A-1387400
2	13.4	89.3	25	7	US-11-121-849-654079
3	13.4	89.3	182314	7	US-11-112-908-45
4	13	86.7	1321	6	US-10-750-185-49780
5	13	86.7	1321	6	US-10-750-623-49780
6	13	86.7	15876	7	US-11-052-554A-660
C 7	12.4	82.7	65	6	US-10-310-914A-7133
C 8	12.4	82.7	65	6	US-10-310-914A-8862
9	12.4	82.7	126	6	US-10-467-657-867
10	12.4	82.7	153	6	US-10-467-657-735
11	12.4	82.7	162	6	US-10-467-657-743
C 12	12.4	82.7	201	6	US-10-995-561-11825
C 13	12.4	82.7	201	6	US-10-995-561-61260
C 14	12.4	82.7	201	6	US-11-052-554A-660
C 15	12.4	82.7	201	6	US-11-212-443-47
C 16	12.4	82.7	558	7	US-11-212-443-49
C 17	12.4	82.7	600	6	US-10-750-185-21924
C 18	12.4	82.7	600	6	US-10-750-623-21924
C 19	12.4	82.7	639	7	US-11-136-527-6844
C 20	12.4	82.7	660	7	US-11-213-443-45
C 21	12.4	82.7	696	6	US-10-467-657-3351
C 22	12.4	82.7	960	6	US-10-750-185-63567
23	12.4	82.7	960	6	US-10-750-623-63567

C 24	12.4	82.7	999	6	US-10-750-185-53582	Sequence 53582, A
C 25	12.4	82.7	999	6	US-10-750-623-53582	Sequence 53582, A
C 26	12.4	82.7	1369	7	US-11-128-061-3328	Sequence 3328, Ap
C 27	12.4	82.7	1400	7	US-11-136-527-1821	Sequence 1821, Ap
C 28	12.4	82.7	1465	6	US-10-750-185-62827	Sequence 62827, A
C 29	12.4	82.7	1465	6	US-10-750-623-62827	Sequence 62827, A
C 30	12.4	82.7	1581	6	US-10-750-185-64370	Sequence 64370, A
C 31	12.4	82.7	1581	6	US-10-750-623-64370	Sequence 64370, A
C 32	12.4	82.7	1604	7	US-11-192-967-7	Sequence 7, Appli
C 33	12.4	82.7	1604	7	US-11-193-715-7	Sequence 7, Appli
C 34	12.4	82.7	1713	6	US-11-136-527-2748	Sequence 2748, Ap
C 35	12.4	82.7	1853	6	US-10-967-527A-1	Sequence 1, Appli
36	12.4	82.7	1938	6	US-10-967-527A-29	Sequence 29, Appl
37	12.4	82.7	1979	6	US-10-967-527A-31	Sequence 31, Appl
38	12.4	82.7	2232	6	US-10-986-501A-79	Sequence 79, Appl
39	12.4	82.7	2879	7	US-11-124-368A-166	Sequence 166, App
40	12.4	82.7	3635	7	US-11-136-527-2101	Sequence 2101, Ap
41	12.4	82.7	4734	7	US-11-136-527-2518	Sequence 2518, Ap
C 42	12.4	82.7	5498	7	US-11-192-967-1	Sequence 1, Appli
C 43	12.4	82.7	5498	7	US-11-193-715-1	Sequence 1, Appli
C 44	12.4	82.7	6719	6	US-10-995-561-456	Sequence 456, App
45	12.4	82.7	9903	7	US-11-052-554A-517	Sequence 517, App

ALIGNMENTS

RESULT 1
US-10-310-914A-1387400
; Sequence 1387400, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
; APPLICANT: Bentwich, Isaac
; TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
; FILE REFERENCE: 06087.0200.CPUS01
; CURRENT APPLICATION NUMBER: US/10/310,914A
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 1388402
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1387400
; LENGTH: 20
; TYPE: RNA
; ORGANISM: Human
US-10-310-914A-1387400

Query Match 89.3%; Score 13.4; DB 6; Length 20;
Best Local Similarity 86.7%; Pred. No. 8.6e+02;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AGCGGGCGGACTCA 15
|||||
Db 2 AGCGGGCGGACTCA 16

RESULT 2
US-11-121-849-654079
; Sequence 654079, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
; APPLICANT: John Palma
; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
; FILE REFERENCE: 3684.1
; CURRENT APPLICATION NUMBER: US/11/121,849
; CURRENT FILING DATE: 2005-05-03
; PRIOR FILING DATE: 2004-05-03
; NUMBER OF SEQ ID NOS: 673904
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 654079
; LENGTH: 25

```
; TYPE: DNA
; ORGANISM: Homo sapien
US-11-121-849-654079

Query Match      89.3%; Score 13.4; DB 7; Length 25;
Best Local Similarity 93.3%; Pred. No. 1.5e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCGGGCGGACTCA 15
   ||||| |||||
Db 6 AGCGGGCTGGACTCA 20

RESULT 3
US-11-112-908-45
; Sequence 45, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; PRIOR FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 45
; TYPE: DNA
; LENGTH: 182314
; ORGANISM: Homo sapiens
US-11-112-908-45

Query Match      89.3%; Score 13.4; DB 7; Length 182314;
Best Local Similarity 93.3%; Pred. No. 1.5e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCGGGCGGACTCA 15
   ||||| |||||
Db 18078 AGCGGGCGGACTCA 18092

RESULT 4
US-10-750-185-49780
; Sequence 49780, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 49780
; LENGTH: 1321
; TYPE: DNA
; ORGANISM: Bovine 19866880919793
```

```
US-10-750-185-49780

Query Match      86.7%; Score 13; DB 6; Length 1321;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCGGGCGGACT 13
   ||||| |||||
Db 169 AGCGGGCGGACT 181

RESULT 5
US-10-750-623-49780
; Sequence 49780, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 49780
; LENGTH: 1321
; TYPE: DNA
; ORGANISM: Bovine 19866880919793
US-10-750-623-49780

Query Match      86.7%; Score 13; DB 6; Length 1321;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCGGGCGGACT 13
   ||||| |||||
Db 169 AGCGGGCGGACT 181

RESULT 6
US-11-052-554A-660
; Sequence 660, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 660
; LENGTH: 15876
; TYPE: DNA
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-660

Query Match      86.7%; Score 13; DB 7; Length 15876;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2006, 16:32:47 ; Search time 54.6 Seconds
(without alignments)
488.341 Million cell updates/sec

Title: US-10-754-446-1

Perfect score: 15

Sequence: 1 agcggcgccgactca 15

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*

- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/PCUS COMB.seq.*
- 7: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
- 8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
- 9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	100.0	15353	3	US-09-949-016-13454
C 2	14	93.3	1950	3	Sequence 13454, A
C 3	14	93.3	10236	3	Sequence 3982, Ap
C 4	14	93.3	10237	3	Sequence 12492, A
C 5	14	93.3	20232	3	Sequence 13859, A
6	13.4	89.3	97	3	Sequence 1197, Ap
C 7	13.4	89.3	616	3	Sequence 21303, A
C 8	13.4	89.3	968	3	Sequence 599, App
C 9	13.4	89.3	981	3	Sequence 7533, Ap
10	13.4	89.3	1281	3	Sequence 99, Appl
11	13.4	89.3	2301	2	Sequence 6674, Ap
C 12	13.4	89.3	2524	3	Sequence 5, Appl
C 13	13.4	89.3	2868	3	Sequence 517, App
C 14	13.4	89.3	2874	3	Sequence 1961, Ap
C 15	13.4	89.3	3048	3	Sequence 1915, Ap
C 16	13.4	89.3	3292	2	Sequence 957, App
C 17	13.4	89.3	3882	3	Sequence 33, Appl
C 18	13.4	89.3	4404	3	Sequence 1005, Ap
C 19	13.4	89.3	44276	3	Sequence 1086, Ap
C 20	13	86.7	387	3	Sequence 17218, A
21	13	86.7	444	3	Sequence 41, Appl
22	13	86.7	471	3	Sequence 11937, A
23	13	86.7	696	3	Sequence 8611, Ap
24	13	86.7	732	3	Sequence 11970, A
					Sequence 11900, A

Sequence 98, Appl
Sequence 12015, A
Sequence 1, Appl
Sequence 259, App
Sequence 259, App
Sequence 15098, A
Sequence 39, Appl
Sequence 39, Appl
Sequence 12957, A
Sequence 11932, A
Sequence 15764, A
Sequence 1, Appl
Sequence 1, Appl
Sequence 69, Appl
Sequence 69, Appl
Sequence 69, Appl
Sequence 69, Appl
Sequence 70, Appl
Sequence 33, Appl
Sequence 35, Appl
Sequence 37, Appl

25 13 86.7 975 3 US-10-152-886-98
c 26 13 86.7 1119 3 US-09-252-391A-12015
c 27 13 86.7 2214 2 US-07-985-458-1
28 13 86.7 4162 3 US-09-453-702B-259
29 13 86.7 4162 3 US-10-114-170-259
c 30 13 86.7 4382 3 US-09-949-016-15098
31 13 86.7 25165 3 US-09-453-702B-39
32 13 86.7 25165 3 US-10-114-170-39
c 33 13 86.7 35471 3 US-09-949-016-12957
c 34 13 86.7 41522 3 US-09-949-016-11932
c 35 13 86.7 41523 3 US-09-949-016-15764
36 13 86.7 58857 3 US-09-477-962-1
37 12.4 82.7 24 3 US-09-507-819-1
38 12.4 82.7 24 3 US-09-507-819-69
39 12.4 82.7 24 3 US-09-507-819-70
40 12.4 82.7 24 3 US-09-641-576-1
41 12.4 82.7 24 3 US-09-641-576-69
42 12.4 82.7 24 3 US-09-641-576-70
43 12.4 82.7 96 3 US-09-507-819-33
44 12.4 82.7 96 3 US-09-507-819-35
45 12.4 82.7 96 3 US-09-507-819-37

ALIGNMENTS

RESULT 1

US-09-949-016-13454
; Sequence 13454, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13454
; LENGTH: 15353
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13454

Query Match 100.0%; Score 15; DB 3; Length 15353;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGCGGCGCGGACTCA 15

Db 6095 AGCGGCGCGGACTCA 6109

RESULT 2

US-09-902-540-3982/c
; Sequence 3982, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(115849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10

```
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 3982
; LENGTH: 1950
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-3982

Query Match          93.3%; Score 14; DB 3; Length 1950;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCGGGCCGGACTC 14
Db 608 AGCGGGCCGGACTC 595

RESULT 3
US-09-949-016-12492/c
; Sequence 12492, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12492
; LENGTH: 10236
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12492

Query Match          93.3%; Score 14; DB 3; Length 10236;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GCGGGCCGGACTCA 15
Db 2595 GCGGGCCGGACTCA 2582

RESULT 4
US-09-949-016-13859/c
; Sequence 13859, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13859
; LENGTH: 10237
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13859

Query Match          93.3%; Score 14; DB 3; Length 10237;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GCGGGCCGGACTCA 15
Db 2595 GCGGGCCGGACTCA 2582

RESULT 5
US-09-902-540-1197/c
; Sequence 1197, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1197
; LENGTH: 20232
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(20232)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1197

Query Match          93.3%; Score 14; DB 3; Length 20232;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCGGGCCGGACTC 14
Db 18886 AGCGGGCCGGACTC 18873

RESULT 6
US-09-513-999C-21303
; Sequence 21303, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: S9.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 21303
; LENGTH: 97
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-21303

Query Match          89.3%; Score 13.4; DB 3; Length 97;
Best Local Similarity 93.3%; Pred. No. 9.5e+02;
```